Utah State University
DigitalCommons@USU

All Graduate Theses and Dissertations | Graduate Studies

12-2021

Soil Health Monitoring and Management in Corn and Soybean Agroecosystems of the Midwestern U.S.

Bradley S. Crookston
Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/etd

Part of the Agronomy and Crop Sciences Commons

Recommended Citation
Crookston, Bradley S., "Soil Health Monitoring and Management in Corn and Soybean Agroecosystems of the Midwestern U.S." (2021). *All Graduate Theses and Dissertations*. 8322.
https://digitalcommons.usu.edu/etd/8322

This Dissertation is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Theses and Dissertations by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.

UtahStateUniversity
MERRILL-CAZIER LIBRARY

SOIL HEALTH MONITORING AND MANAGEMENT IN CORN AND SOYBEAN AGROECOSYSTEMS OF THE MIDWESTERN U.S.

by

Bradley S. Crookston

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Soil Science

Approved:

Matt A. Yost, Ph.D. Major Professor	Jeanette Norton, Ph.D. Committee Member
Grant Cardon, Ph.D. Committee Member	John Stevens, Ph.D. Committee Member
Kristen Veum, Ph.D. Committee Member	D. Richard Cutler, Ph.D. Interim Vice Provost of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2021

Copyright © Bradley S. Crookston 2021

All Rights Reserved

ABSTRACT

Soil Health Monitoring and Management in Corn and Soybean

Agroecosystems of the Midwestern U.S.

by

Bradley S. Crookston, Doctor of Philosophy

Utah State University, 2021

Major Professor: Dr. Matt A. Yost
Department: Plant, Soils and Climate

Monitoring and managing soil health at the farm level relies on databases that characterize the relationships between soil health indicators and soil services like crop yield. Data were collected from over 100 farms across the Midwestern US that were members of an on-farm participatory research network called the Soil Health Partnership, which operated from 2014 to 2021. Twelve soil health indicators used in three common soil health assessments were analyzed along with corn (*Zea mays* L.) and soybean (*Glycine max* L.) yield to explore relationships among soil health indicators, scores, and crop yield. Three studies were conducted to 1) evaluate the influence of temporal and spatial variation in soil health indicators on yield variability; 2) determine the correlation strength among soil health assessment scores and the number of site-years that scores are correlated with yield; and 3) determine the effects of cover crops on soil health indicators, scores, and yield.

Multiple regression revealed that indicator variability accounted for relatively slight variation in yield. Additionally, simple regression showed that yield is more correlated with indicator scores at the site level rather than within individual site years. Finally, analysis of covariance with repeated measures demonstrated that the effect of one to four years of cover crops is minimal on soil health indicators; only 96-hr carbon mineralization was affected by cover crops. These results demonstrate that cash crop yield is an unclear metric of soil health. Overall, these results may suggest to growers a whole-of-ecosystem approach to monitoring soil health and that soil health measurements be collected before beginning a new conservation management plan, then every two to four years to allow time for soil health improvement.

(146 pages)

PUBLIC ABSTRACT

Soil Health Monitoring and Management in Corn and Soybean

Agroecosystems of the Midwestern U.S.

Bradley S. Crookston

Soil health is a concept and condition of the soil where measurable soil properties represent the capacity of a soil fulfilling its intended use, such as producing crops, without constraint to its agro-ecological quality. Soil health assessments are used to estimate the health of a soil by assessing soil biological, chemical, and physical attributes, called soil health indicators, and scoring them on a scale, usually 0 to 100, to guide soil and crop management. However, there are few large-scale analyses of soil health assessment scores and their relationships with crop yield. Understanding how soil health assessments relate to crop yield can support soil health practitioners and growers in making decisions that can direct efforts to improve soil health monitoring and management.

The Soil Health Partnership (SHP) was a sizeable farmer-led network of on-farm trials assessing soil health throughout the Midwestern US. The on-farm data was used to explore the relationship between soil health and crop yield in three ways. First, how variability in soil health affects variability in yield. Second, the strength of the relationships between soil health assessment scores and crop yield. And third, the effects of conservation management on soil health indicators, scores, and yield.

These analyses found that soil health indicator variation in time accounted for relatively little variability in corn and soybean yield over a two-to-four-year timespan at the SHP sites. Second, soil health scores of individual indicators or composite scores were not often correlated with crop yield on a site-to-site basis. This might suggest to soil health researchers and growers that other soil health outcomes, such as field runoff water quality, be measured to determine how soil health is improving additional soil ecosystem services. Third, the on-farm soil health trials revealed that few soil health indicators were affected by cover crops within a short one to four years of treatment timespan. Overall, these results suggest to growers that a whole-of-ecosystem approach be taken to monitoring soil health and that soil health measurements be taken before beginning a new conservation management plan, then every two to four years to allow time for soil health improvement.

ACKNOWLEDGMENTS

I would like to thank Dr. Matt Yost for making the time to guide me as I begin the journey toward doctoral expertise. I would especially like to thank my committee members, Drs. Kristen Veum, Grant Cardon, Jeanette Norton, and John Stevens, for their kindness, patience, and support throughout my time as a student. I would also like to thank Maria Bowman for providing feedback and mentorship that dared me to probe my assumptions and challenge my conclusions. I also extend special thanks to the former staff of the Soil Health Partnership, especially Jack Cornell, who provided support and effort that made this dissertation possible.

Finally, to my wife, family, friends, and fellow students, I give thanks for their encouragement and moral support. I could not have done it without all of you.

Bradley S. Crookston

CONTENTS

LIST OF TABLES

LIST OF TABLES, CONT.

LIST OF FIGURES

LIST OF FIGURES, CONT.

CHAPTER 1

INTRODUCTION

Soil health is a concept and condition of the soil where measurable abiotic and biotic soil properties represent processes relevant to a soil fulfilling its functional intended use without constraint to its environmental quality (Andrews et al., 2004; Doran & Zeiss, 2000; Haney et al., 2018; Moebius-Clune, et al., 2016). Soil health is quantified by biological, chemical, and physical soil health indicators that are intended to be representative of ecosystem services and functions; responsive to soil and crop management changes without merely echoing natural annual cycles in weather or management; analytically affordable and timely; and valuable in management decision-making by indicating constraints to soil health (Lehmann et al., 2020). However, interpreting an assessment of soil health relies on databases regionally calibrated to edaphic and climatic factors, so indicator values are comparable to similar soils (Fine et al., 2017; Nunes et al., 2021). Furthermore, soil health interpretive frameworks, otherwise known as soil health assessments, were developed to estimate soil health by translating an observed value from a property analyzed in a soil sample into a unitless score to interpret inherent potential for soil health and dynamic responses to management practices (Stott, 2019). Some scoring methods are based on site-specific conditions (climate and soil type) and crop factors (Andrews et al., 2004), while others are based on regional and soil textural categories (Fine et al., 2017), soil type and climate peer groups (Nunes et al., 2021), or soil property thresholds (Haney et al., 2018). Understanding relationships among the types of assessments and their relationship to soil health outcomes, such as crop yield, is crucial to utilizing and interpreting soil health assessments.

As a multifaceted concept, soil health intersects with soil security, a framework that elevates soils as an integral component of global environmental sustainability goals crucial to addressing existential challenges for humanity. These challenges include water, food, and energy security, climate change mitigation, biodiversity protection, and maintaining ecosystem services (Amundson et al., 2015; McBratney et al., 2014). The soil security concept seeks to secure soils in the same sense that food and water can be secured to ensure availability, quality, and use for humanity (McBratney et al., 2014). Soil health assessment is interconnected with the five dimensions of soil security: capability, condition, capital, connectivity, and codification, to assess what soil can do while addressing how soil can continue to function and garner human interest in the soil through value and policy. Consequently, promoting the adoption of soil health assessment and use of soil health-promoting practices also supports growers and agricultural practitioners in addressing societal, environmental concerns.

Soil health-promoting practices are agricultural soil and crop management strategies intended to curb soil loss and degradation while improving soil health. Healthy soil has sufficient nutrient supply, biological activity, and good soil structure, for example (Magdoff & van Es, 2009), and ultimately, is resilient to degradation. Among many soil health-promoting practices, reducing soil tillage and maintaining year-round live roots and soil cover using cover crops have been shown to improve soil health (Magdoff & van Es, 2009). The many benefits of soil health-promoting practices in production agriculture have been explored, and in many instances, have investigated soil health indicators and assessment scores (Marcillo & Miguez, 2017; Nunes et al., 2020;

Yang et al., 2020). However, there is still a need to show the effects of soil health-promoting practices in on-farm participatory research where growers are directly involved in implementing and managing a soil health experiment on their farm. The Soil Health Partnership (SHP) was a program led by growers of the National Corn Growers Association and supported on-farm soil health research from 2014 to 2021. The SHP brought together collaborators from federal agencies, private companies, farmer groups, universities, and environmental groups to promote soil health practices for economic and ecological benefits. The SHP worked with growers throughout much of the Midwestern US by establishing randomized and replicated strip trial comparisons of a grower's historical management versus a soil health-promoting practice, typically cover crops or reduced tillage. Growers then provided general information about their agronomic practices and crop yield, and soil samples are regularly collected and analyzed for a comprehensive suite of common soil health indicators. The SHP dataset was used in this manuscript to explore the overarching question, how well do soil health assessments relate to corn and soybean yield? in three ways. First, by evaluating temporal and spatial variation in soil health indicators and grain yield. Second, by determining the strength of relationships among three common soil health assessments and yield. Third, determining the effect of cover crops on soil health indicators, scores, and yield. This work aims to support soil health practitioners and growers in their efforts to interpret, monitor, and manage soil health for crop productivity.

REFERENCES

Amundson, R., Berhe, A. A., Hopmans, J. W., Olson, C., Sztein, A. E., & Sparks, D. L. (2015). Soil and human security in the 21st century. *Science*, *348*(6235). https://doi.org/10.1126/science.1261071

Andrews, S. S., Karlen, D. L., & Cambardella, C. A. (2004). The Soil Management Assessment Framework. *Soil Science Society of America Journal*, *68*(6), 1945–1962. https://doi.org/10.2136/sssaj2004.1945

Doran, J. W., & Zeiss, M. R. (2000). Soil health and sustainability: managing the biotic component of soil quality. *Applied Soil Ecology*, *15*(1), 3–11. https://doi.org/10.1016/S0929-1393(00)00067-6

Fine, A. K., van Es, H. M., & Schindelbeck, R. R. (2017). Statistics, Scoring Functions, and Regional Analysis of a Comprehensive Soil Health Database. *Soil Science Society of America Journal*, *81*(3), 589–601. https://doi.org/10.2136/sssaj2016.09.0286

Haney, R. L., Haney, E. B., Smith, D. R., Harmel, R. D., & White, M. J. (2018). The soil health tool—Theory and initial broad-scale application. *Applied Soil Ecology*, *125*, 162–168. https://doi.org/10.1016/j.apsoil.2017.07.035

Lehmann, J., Bossio, D. A., Knabner, I. K.-, & Rillig, M. C. (2020). The Concept and Future Prospects of Soil Health. *Nature Reviews Earth & Environment*, *1*, 544–553. https://doi.org/10.1038/s43017-020-0080-8

Magdoff, F., & van Es, H. (2009). *Building soils for better crops: sustainable soil management* (3rd ed.). Sustainable Agriculture Research and Education (SARE).

Marcillo, G. S., & Miguez, F. E. (2017). Corn yield response to winter cover crops: An updated meta-analysis. *Journal of Soil and Water Conservation*, *72*(3), 226–239. https://doi.org/10.2489/jswc.72.3.226

McBratney, A. B., Field, D. J., & Koch, A. (2014). The dimensions of soil security. *Geoderma*, *213*, 203–213. https://doi.org/10.1016/j.geoderma.2013.08.013

Moebius-Clune, B. N., D.J. Moebius-Clune, B.K. Gugino, O.J. Idowu, R.R. Schindelbeck, A.J. Ristow, H.M. van Es, J.E. Thies, H.A. Shayler, M.B. McBride, K.S.M Kurtz, D.W. Wolfe, & G.S. Abawi. (2017). *Comprehensive Assessment of Soil Health – The Cornell Framework* (3.2). Cornell University. http://soilhealth.cals.cornell.edu/training-manual/

Nunes, M. R., Karlen, D. L., Veum, K. S., & Moorman, T. B. (2020). A SMAF Assessment of U.S. Tillage and Crop Management Strategies. *Environmental and Sustainability Indicators*, *8*. https://doi.org/10.1016/j.indic.2020.100072

Nunes, M. R., Veum, K. S., Parker, P. A., Holan, S. H., Karlen, D. L., Amsili, J. P., Es, H. M., Wills, S. A., Seybold, C. A., & Moorman, T. B. (2021). The soil health assessment protocol and evaluation applied to soil organic C. *Soil Sci. Soc. Am. J.*, *85*, 1196–1213. https://doi.org/10.1002/saj2.20244

Stott, D. E. (2019). *Recommended Soil Health Indicators and Associated Laboratory Procedures. Soil Health Technical Note No. 450-03. U.S. Department of Agriculture, Natural Resources Conservation Service.* https://www.nrcs.usda.gov/wps/PA_NRCSConsumption/download?cid=nrcseprd1420229&ext=pdf

Yang, T., Siddique, K. H. M., & Liu, K. (2020). Cropping systems in agriculture and their impact on soil health-A review. *Global Ecology and Conservation*, *23*, e01118. https://doi.org/10.1016/j.gecco.2020.e01118

CHAPTER 2

SOIL HEALTH SPATIAL-TEMPORAL VARIATION INFLUENCE SOIL SECURITY ON MIDWESTERN, U.S. FARMS[1]

ABSTRACT

Soil security is a multifaceted framework that considers soil as an integral part of addressing societal concerns towards global environmental challenges. Soil health assessments are tools that can be used to integrate knowledge about and social interest in soil resource sustainability. Appropriate interpretation of soil health assessments requires robust databases of soil properties and their variation across large regional areas. This analysis explored field-scale spatial and temporal variation in 16 soil health indicators used in common soil health assessments at Soil Health Partnership (SHP) locations throughout the Midwestern U.S. from 2014–2019. Relationships among management, environment, and measured soil properties were examined using various combinations of correlation, principal component analysis (PCA), and multiple regression. Specifically, variability was evaluated using 1) the temporal average of indicator lab test values, 2) the temporal and spatial coefficient of variation (CV), and 3) corn (*Zea mays* L.) and soybean (*Glycine max* L.) yield variation. Solvita® had the highest spatial and temporal CV, while organic matter (OM), autoclaved citrate extractable protein (ACE), and pH had the lowest

---

[1] Coauthors: Matt A. Yost, Maria Bowman, Kristen Veum, Grant Cardon, Jeanette Norton; An article published in *Soil Security*, March 2021, https://doi.org/10.1016/j.soisec.2021.100005.

spatial and temporal CV values. The PCA analysis identified climate, soil texture, organic C and N pools, and soil water availability as factors that accounted for variation in soil health indicator values. Multiple regression showed that climate variables and field conditions strongly affect corn and soybean yield variation. Solvita, OM, and available water content improved corn and soybean yield variation estimates. These results show that considering spatial and temporal variation when monitoring soil health changes may improve soil health assessment interpretation.

Keywords: Soil Security; Soil health; Indicators; Variation

Abbreviations: ACE, autoclaved citrate extractable soil protein; ActC, active carbon; AggStabl, aggregate stability; AWC, available water content; AWDR LT avg, abundant and well distributed rainfall long-term average; AWDR TCV, abundant and well-distributed rainfall temporal coefficient of variation; GDD LT avg, growing degree day long-term average; GDD TCV, growing degree day temporal coefficient of variation; OM, organic matter; Resp, microbial respiration (4-day incubation); SHInd, soil health indicator; WEOC, water-extractable organic carbon; WEON, water-extractable organic nitrogen.

1. Introduction

Soil security is a framework that elevates soils as an integral component of global environmental sustainability goals that are crucial to addressing existential challenges for humanity. These challenges include water, food and energy security, climate change mitigation, biodiversity protection, and maintaining ecosystem services (Amundson et al., 2015; McBratney et al., 2014). As a multifaceted concept, soil security is composed of five major dimensions related to securing soils in the same sense that food and water can be secured to ensure availability, quality, and use for humanity (McBratney et al., 2014). Those five dimensions are capability, condition, capital, connectivity, and codification. Capability is asking the question "what can a soil do?" while condition is addressing "can this soil continue to do this?" Connectivity, capital, and codification are related to how humans interact with and value soil resources and implement policy for soil management. Soil health assessments, such as the Comprehensive Assessment of Soil Health (CASH), Soil Management Assessment Framework (SMAF), and the Haney Soil Health Tool (HSHT), utilize soil condition (health) indicators – measurable abiotic and biotic soil properties that represent processes relevant to a soil fulfilling its functional capability without constraint to its condition (Andrews et al., 2004; Doran & Zeiss, 2000; Haney et al., 2018; Moebius-Clune et al., 2017). Interpretable and suitable soil health indicators are representative of ecosystem services and functions; responsive to management changes without merely echoing natural annual cycles; analytically affordable and timely; and useful in management decision-making (Lehmann et al., 2020). Interpreting a soil health assessment relies on databases regionally calibrated to edaphic and climatic factors so indicator values are comparable to similar soils (Fine et al., 2017; Nunes et al., 2021). However, measures of temporal and spatial variation are not typically included with indicator values in regional databases (Baveye et al., 2016). Thus, reporting uncertainty is necessary to validate soil condition benchmark values while monitoring for soil health changes at the field or regional scale to inform soil codification and valuation of cultivated natural capital.

1.1. Dimensions of the Soil Security Framework

Soil is connected to major societal challenges through seven soil functions: (i) biomass production, (ii) storing, filtering and transforming of nutrients and water, (iii) biodiversity pool, (iv) physical and cultural environment, (v) raw material source, (vi) C pool and cycling, and (vii) geological and cultural heritage archive (McBratney et al., 2014). However, soils inherently differ in their potential capacity to fulfill these functions based on climate, duration of soil development, organismic influence on the soil (including humans, vegetation, and meso/microorganisms), soil source (parent) material, and topographical relief (Jenny, 1994) which are all accounted for in soil taxonomy. Soil capability is concerned with a soil's potential to fulfill soil functions as defined by its "natural" state in relation to features defined in its taxonomy (McBratney et al., 2014). The SMAF provides a convenient classification of soil taxonomy suborders based on their relative potential for C sequestration and C cycling (Stott et al., 2010), relevant for cultivated natural capital production, that can be used for a reference frame for discussing soil security dynamics. Soil condition is an evaluation of a soil's performance compared to the "natural" potential (McBratney et al., 2019).

Human interaction and soil management are represented in the soil capital, connectivity, and codification dimensions. Soil capital refers to the monetization and valuation of soil stocks – those that are renewable, nonrenewable, replenishable, and cultivated (McBratney et al., 2014). McBratney et al. (2014) noted that models are needed to quantify the portion of cultivated natural capital, as in crop or agroforestry production, that can be attributed to soil capability and condition to better evaluate the efficacy of new management adaptations. Social interaction with soil is represented in the connectivity dimension and is in part concerned with soil as a public good. Connectivity also represents the social value of soil as healthy soils contribute to human well-being among many other social goods (Friedrichsen et al., 2021). Thus, soil health assessment represents a connectivity linkage between soil as a stock of natural cultivated capital with inherent capabilities and conditions based on its current use by land "stewards." Finally, to ensure secure soils, initiatives and public policies that design, implement and evaluate dimensions of soil security are needed to draw interest from appropriate stakeholders, e.g., growers, farm financial lenders, and policymakers. While organizations like the Soil Health Institute, Soil Health Partnership (SHP), and The Nature Conservancy (TNC), among many others, are making great strides in promoting the soil health concept, additional work can be done to raise awareness toward soil's multidimensionality as characterized in the soil security framework.

1.2. Quantifying Soil Health to Improve Assessment Interpretation

Recent studies have quantified soil health indicator values in differing environments and scales under diverse agronomic systems (Fine et al., 2017; Haney et al., 2018; Svoray et al., 2015; Valani et al., 2020); however, few regional databases presently exist with descriptions of soil health indicator temporal and spatial variation and their associated with crop yield variation. These measures of variation are needed to improve practitioners' criteria for soil health indicator interpretability and suitability. They may also improve indicator selection strategies, sampling regimes, and determining soil health progress following management adaptations to ensure future sustainable agroecosystems (Doran & Zeiss, 2000; Rinot et al., 2019).

Several geographically large projects focusing on soil health in various countries are currently underway. The Cornell Atkinson Center for Sustainability, TNC, and the China Agriculture University are collaborating on a project located in China to promote sustainable land management by calibrating CASH soil health indicators to agricultural soils in China (Lawrence & Friedlander, 2020). The North American Project to Evaluate Soil Health Measurements, administered by the Soil Health Institute, collects soil health data from long-term agriculture research locations and working farms to help establish standardized measurements for soil health (Norris et al., 2020). Additionally, the Cornell University Soil Health Laboratory has developed an extensive database of soil health indicators predominately from the Northeastern U.S., but increasingly from other regions such as the Midwest and central Atlantic coast (Fine et al., 2017). Although databases like these are necessary for soil health assessment, there is not yet a Midwestern U.S. regional database of common indicators from working farms that also summarizes temporal and spatial variation in connection with crop yield variation.

The Soil Health Partnership is a grower-led initiative of the National Corn Growers Association that was established in 2014. The SHP has brought together collaborators from federal agencies, private companies, farmer groups, universities, and environmental groups to promote soil health practices for economic and environmental benefits. The SHP works directly with growers throughout much of the Midwestern US by establishing on-farm randomized and replicated strip trial comparisons of grower's historical management versus a soil health-promoting practice, typically cover crops or reduced tillage. Growers provide general information about their agronomic practices and crop yield, and soil samples are regularly collected and analyzed for a comprehensive suite of common soil health indicators.

Uniquely, this dataset represents working farms from a large geographic area, allowing research results from the SHP database to have broad generalizability to farms within the region. This study's objectives were to summarize soil health indictor values and temporal and spatial coefficient of variation (CV) across SHP locations in the Midwestern U.S. and describe the relationship of soil health indicator temporal variation with corn and soybean yield variation. Ultimately, this research aims to help fill gaps in soil health databases that support assessment interpretation by making available summaries of soil conditions at working farms.

2. Methods

2.1. Data Collection from On-Farm Evaluations of Soil Health

Soil Health Partnership data collection procedures are standardized for all locations. For detailed information regarding the establishment of the on-farm soil health strip trials, see https://www.soilhealthpartnership.org. On-farm trials began in 2014 at 14 fields across Illinois, Indiana, Iowa, Nebraska, and Ohio. Additional locations joined the network in subsequent years in Florida, Kentucky, Maryland, Michigan, Minnesota, Missouri, North Dakota, South Dakota, Tennessee, and Wisconsin.

On-farm soil health evaluation trials included randomized strips of the grower's historical management as a control treatment and a soil health treatment, typically a cover crop application or reduced tillage, each replicated four times. Samples for nutrient and soil health analyses were collected from SHP locations in the spring before planting, using predetermined geolocated 0.4 ha grid sampling points across strips that range from 0.4–4 ha. Soil cores are randomly sampled from within a 4.5 m radius of the geolocated point. Soil samples for nutrient analyses were collected to 15 cm and then separated into 0–5 cm and 5–15 cm sections. Samples for soil health analyses (e.g., soil respiration or protein) remained in one segment (0–15 cm), composited across a strip, and placed in coolers with ice packs for expedited shipping to soil analysis laboratories. Soil analyses at Ward Laboratories, Inc. (Kearney, NE) included soil organic matter (OM) loss-on-ignition, pH, phosphorus (P), potassium (K), sulfur (S), calcium (Ca), and magnesium (Mg). Soil health samples in coolers were sent to Cornell University Soil Health Laboratory (Ithaca, NY) for analyses of soil active carbon (ActC), 96-hour respiration assay (Resp (4 d)), autoclaved citrate extractable protein index (ACE), wet aggregate stability (AggStabl), and available water capacity (AWC). The USDA-Agriculture Research Service Grassland, Soil and Water Research Laboratory in Temple, TX received additional samples in coolers from 2014 to 2018 for the Haney Soil Health Tool (HSHT) suite analysis (Solvita CO2-Burst, water-extractable organic C (WEOC), and water-extractable organic N (WEON; Haney et al., 2018). Ward Labs completed the HSHT analyses during 2019. Soil analysis and procedures are summarized in the following section.

Sampling for soil nutrient analyses at SHP locations took place approximately annually between 2014 and 2019. Samples for soil health analyses were collected at SHP locations approximately every-other-year beginning in 2015. Corn and soybean yield data were collected annually from combine-mounted yield monitoring systems with corresponding global positioning system signal locations. Yield data quality assurance followed the procedure outlined by the Iowa Soybean Association (Kyveryga et al., 2018) and were averaged across each strip.

2.2. Soil Health Indicator Laboratory Analyses

The suite of soil health indicators utilized in this analysis were those associated with the most common soil health assessments in the U.S. and most likely to be utilized by growers and research practitioners in the Midwest, namely, the CASH, SMAF, and HSHT. Details of the current laboratory methods, protocols, and procedures used at the time of analysis for these soil samples are available from Ward Laboratories, Inc. (Ward Laboratories, 2020) Cornell University Soil Health Laboratory (Schindelbeck et al., 2016), and the USDA-ARS Grassland, Soil and Water Research Laboratory (Haney, 2020). A summary of the indicators, a description of their general purpose, and analysis methods are available in Table 1. Although some of the indicators measure similar soil properties (e.g., Resp (4 d) and Solvita measure CO^2 respired by soil microbial communities), these are quantified by different methods and are utilized in different soil health assessments (i.e., CASH, and HSHT). Consequently, it is important to assess many indicators to compare their temporal and spatial variability.

2.3. Dataset Preparation

Weighted averages were calculated for the two-depth nutrient soil samples. Results were reported for the 0–15 cm depth layer to ensure nutrient and soil health test results' comparability. Results were then averaged within a strip for each nutrient and soil health test. The dataset included control strips at 96 SHP locations covering fields added to the network between 2014 and 2018. Of the 96 locations, ten had one year of soil health data, 65 locations had two years, 15 had three years, and six had four years of data. For spatial analyses, all sites with soil nutrient and health data were included. Temporal analyses were conducted on locations with at least two years of soil nutrient and health data. For corn and soybean yield temporal analyses, the dataset was filtered to locations with yield data for the same crop across multiple years. Also, not all soil health indicators were measured at all sites in all sampling years. Therefore, some indicators had a larger number of observations than others in the spatial and temporal analyses.

2.3.1. Climate Data

Daily precipitation and daily maximum and minimum air temperature were collected using the Daymet Single Pixel Extraction Tool (Oakridge National Laboratory, 2020) for the latitude and longitude corresponding to each site. Pixels in the Daymet database represent 1 km2 of interpolated data. Cumulative growing degree days (GDD) were calculated using a ten-degree Celsius base temperature and 30°C as the crop maximum (North Dakota Agriculture Weather Network, 2019). Abundant and well-distributed rainfall (AWDR), a diversity measure calculated from daily rainfall, is used to describe temporal precipitation availability and is calculated using cumulative and daily precipitation in a given period multiplied by the Shannon Diversity Index (Tremblay et al., 2012). A low AWDR index implies unevenness throughout a period in conjunction with low rainfall, whereas a high value represents a more evenly distributed and a greater amount of rainfall. Growing degree days and AWDR were calculated for the period from 1 Apr to 31 Oct representing a typical growing season. Official long-term climate normal values were difficult to obtain for every SHP site; therefore, a growing season long-term average (LT avg) was calculated for weather parameters from 1983 and 2013. The long-term coefficient of variation (TCV) was calculated for AWDR and GDD to represent variability in weather conditions at each site.

Throughout the study region, at approximately 70% of the SHP locations, the long-term average accumulated GDD (GDD long-term average (31 years)) in a growing season (1 Apr to 31 Oct; 1983–2013) was 1600–2000 degree days (°C) with a temporal CV across years of 5.5–7%. Across all locations, the GDD long-term average ranged from approximately 1300 to 2200 degree days. Additionally, 70% of the SHP locations were rated at >460 AWDR, and a temporal CV of 23–28% in approximately 50% of locations (see Supplementary Fig. 2−S1). A strong negative correlation (r = -0.92) was observed between GDD long-term average and GDD temporal CV (Supplementary Table 2−S1), indicating that locations with high GDD have lower temporal variation within a season than locations with lower seasonal GDD accumulation. In AWDR, there was no correlation between the long-term average and temporal CV.

2.3.2. Soil Taxonomy and Management Data

The predominant soil texture classes represented by the 96 SHP locations were silt loam (52% of locations), silty clay loam (17%), loam (13%), sandy loam (9%), clay loam (6%), and loamy sand (3%). Soil suborder taxonomy was identified using the USDA-NRCS Web Soil Survey (Soil Survey Staff, 2019) area of interest tool for each location and grouped according to the same typological system used in the Soil Management Assessment Framework (SMAF); namely, suborders are classified by climate/moisture regimes and inherent OM levels corresponding to their potential to sequester C (Andrews et al., 2004; Stott et al., 2010) (Table 2). For the SMAF soil suborder class 2, soils in this dataset were members of the Alboll, Aquoll, Udoll, and Ustoll of the Mollisols order. Soils in class 3 were members of the Aqualfs, Udalfs, and Ustalfs of the Alfisols order. Soil characteristics within suborder class 2 have a greater potential to sequester C relative to suborder class 3 (Stott et al., 2010). Sixty percent of the SHP sites were in soil suborder class 2, while the remaining 40% were in class 3. Field topographical slope was identified by the dominant estimated slope from the Web Soil Survey for each location and classified into slope groups for analysis (Table 2).

Each SHP location had both treated and non-treated strip plots. For all analyses, only data from the non-treated strips were used. The purpose of this was to examine inherent variability in soil health indicators without interference from treatments. Evaluation of treatment impacts on soil health indicators and assessment scores at SHP locations are forthcoming. The tillage method in the non-treated strips was the management variable used for this analysis. Among the locations in this dataset, less than 25% continuously practice a conventional (disk plow) form of tillage. The remainder used a form of reduced (strip or vertical) or no-tillage. All sites in the dataset used annual rotations of corn and soybean.

2.4. Statistical Analyses

Calculation of spatial and temporal CV utilized the MEANS procedure of SAS (SAS Institute Inc., 2020). Spatial CV was calculated across strips for each year of data at each location, and temporal CV was calculated for each strip at each location across years. Descriptive statistics box plots for soil properties and crop yield used the location average. Analysis of variance for soil health indicator temporal and spatial CV used Fisher's least significant difference (LSD) in the agricolae package (de Mendiburu, 2020) of the R programing language version 4.0.2 hosted in R Studio. The CORR procedure of SAS was used to calculate the Pearson correlation coefficients. Principal component analysis (PCA) was completed using the PRINCOMP procedure of SAS. Variables in the PCA were evaluated for contributions to dataset variation using principal component (PC) eigenvector loading by identifying the most highly weighted variables in a PC. In the PCA, a weight is given to each variable in each PC based on the correlation of each variable with other variables and their contribution to variability in the PC; each PC represents a proportion of the total variation in the dataset (James et al., 2017; Yeater et al., 2015). Interpretation of each PC utilizes the high weights (high relative absolute value) given to variables within a PC. A meaningful structure in the multivariate dataset can be derived using the dominant variable's weight to "label" the PC as a composite of the variables (James et al., 2017; Yeater et al., 2015). A PC weight of 0.3 (absolute value) was used as a baseline threshold. Two PCA's were conducted: first, with the soil health indicator temporal average and second, with the indicator temporal CV values, each combined with location climate, field, and tillage management variables. These two multivariate datasets were used to investigate temporal variation relationships among the soil health indicators and location condition variables.

Analysis of temporal variation utilized best subsets multiple regression to help explain, rather than predict, temporal trends (temporal average) and variation (temporal CV) in corn and soybean yield. The best subsets regression methodology evaluated all possible combinations of explanatory variables in the regression model to identify a minimum set that best estimates the dependent variable. The temporal average and temporal CV values were calculated for each strip at each location across years. There were 25 explanatory variables: 16 soil health indicators combined with silt and clay concentrations, four climate variables (AWDR/GDD long-term average and CV), soil suborder (two levels), field slope (three levels), and tillage method (four levels) as location factors (Table 1). Four combinations of dependent and explanatory variables resulted in eight models, four each for corn and soybean yield (Fig. 2−1). The four combinations of models were used to explore the complex interactions of measures of temporal variation in soil health indicators, location factors, and crop yield.

To perform the regression analysis, the data were randomly split to create a model training set and a model validation set (70 and 30% of the original set, respectively). Dependent variable (corn and soybean yield temporal average and temporal CV) assumptions of normality of residuals, constant error variance (homoscedasticity), and independence of errors were tested using a custom macro in SAS. First, the assumption of normality of residuals was tested to ensure the distribution of errors was normal using the correlation between the observed errors and the normal-expected errors. The criterion for rejection was a correlation coefficient below 0.983, based on the sample size (n = 175) and a significant alpha level of 0.05 (Kutner et al., 2004). Second, a constant variance was evaluated using the Brown-Forsythe significance test for non-constant variance, where the significant alpha level was 0.05 for the null hypothesis of a constant variance. Third, a sequence plot was used to identify any violations of independence of errors. No violations of the assumptions were found for any dependent variable.

The model statement option *selection = AdjRSq Cp AIC SBC* of the REG procedure, of SAS generated all possible explanatory variable subsets. The criterion for selection identified the Mallow's C_p (C*p*) value that was approximately the number of explanatory variables (in the best subset) plus 1, the lowest Akaike's information criteria (AIC) and Schwartz's Bayesian criterion (SBC), and the highest Adjusted R^2 value (AdjRSq) to accomplish dual goals of maximizing variation explained and minimizing risk of overfitting. Models were sorted by each criterion to identify a model that best fits these requirements. In the case of a tie, the parsimonious model was chosen. Finally, the model intercept and coefficients were validated on the reserved data using the SCORE procedure in SAS. Evaluation of model over-fit utilized the ratio of the mean square prediction error (MSPE) of the test dataset to the mean square error (MSE) of the training set; a ratio of more than 10 indicated over-fitting. The MSPE was defined as:

$$MSPE = \frac{\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}{n}$$

where Y_i is the observed value and $\hat{Y}_i$ is the predicted value in the validation dataset. Root mean square error (RMSE), the square root of the MSPE, was also used to measure model error performance (Chai & Draxler, 2014).

To evaluate the impact that soil health indicators and location condition variables had on corn and soybean yield variation, centered and scaled regression coefficients were utilized to standardize units of the explanatory variables, so they were comparable within and among models. Specifically, the original regression coefficients were multiplied by the ratio of the dependent variable's standard deviation to the explanatory variable's standard deviation. The standardized regression coefficients estimate the number of standard deviations the dependent variable will change for one standard deviation change in an explanatory variable, all others being equal (Wicklin, 2018). The STB option in the model statement of the REG procedure of SAS produced the standardized regression coefficients for each model.

3. Results and discussion

3.1. Soil Capability and Condition: Soil Health Indicator Values and Variation

3.1.1. Range of Values for Soil Health Indicators and Crop Yield

Sixteen soil properties used as indicators of soil health that are measured in three major soil health assessments (CASH, SMAF, and HSHT) were evaluated across 96 locations in the SHP network. Mean values observed at SHP sites (Fig. 2−2) were similar to those reported by other studies in the U.S. For example, values for OM, pH, ACE, and Resp (4 d) reported by Fine et al. (2017) from approximately 900 soil samples throughout Midwestern states, were less than 20% different (absolute value) from SHP average values. Furthermore, mean AggStabl and AWC were 30 and 25% different (absolute value) from values observed by Fine et al. (2017), respectively. Mean OM and Solvita values at SHP sites were 3 and 21% greater, respectively, than from 17 farms in eight Midwestern states reported by Yost et al. (2018). Another study in Ohio found that variability across three locations had a range of 2–23% (CV) for pH, ACE, ActC, OM, Resp (1 d) in increasing order (Hurisso et al., 2018). Values for WEOC, WEON, and the WEOC:WEON ratio (hereafter C:N) from a clay loam soil in San Joaquin Valley, California, were also similar to those observed at SHP locations (9–17% different, absolute value); however, average Solvita values at SHP locations were five times greater than those in the semiarid climate of central California (Mitchell et al., 2017). At the SHP locations, corn yield ranged from 6 to 15 Mg ha-1, and soybean yield ranged from 2 to 6 Mg ha-1 (Supplemental Fig. 2−3 A).

Soil suborder classes 2 and 3 (Table 2−1) often influenced the range of soil biological, chemical, physical property values, and crop yield (Fig. 2−2). The suborder classes (Table 2−2) are differentiated by soil forming factors, such as climate (moisture and temperature that differ across geography), that led to greater OM in suborder class 2 compared to class 3 (Fig. 2−2). Increased biological properties in suborder class 2, e.g., ACE and C:N, may be related to increased clay content that tend to protect soil microbial biomass (Six et al., 2006) in those suborders relative to class 3. Additionally, Stott et al. (2010) found that microbial activity is also related to soil taxonomy. Corn yield was also greater in suborder class 2, which also illustrated how soil capability and potential for higher soil condition can influence soil natural capital in certain crops (Fig. 2−2). Lal (2016) described that inherent soil capability (Lal used the term soil quality) is related to ecosystem services and the outcomes of soil condition (health) such as C pool dynamics, soil structure, water retention and aeration, nutrient cycles, and gaseous emissions that moderate atmospheric CO_2.

3.1.2. Variation in Soil Health Indicators

Temporal and spatial CV was used to assess variation in soil health indicators at the SHP locations. A low negative correlation between spatial CV and field size for eight of the 16 indicators (mean adjR^2 = 0.03, $p < 0.05$), six of which were biological soil properties, indicated that plot size did not have a substantial effect on spatial CV (data not shown). The ANOVA of spatial CV at 94-96 locations was conducted among the biological, chemical, and physical indicators (Fig. 2−3 A). Compared to other biological indicators, Solvita was 13 percentage points greater than Resp (4 d) and approximately 20 points greater than OM, C:N, and ACE. In the chemical category, pH had the lowest, and P had the highest spatial CV, yet the spread between Ca, S, Mg, and K was approximately 4 CV percentage points. Aggregate stability spatial CV was roughly 12 percentage points higher than AWC. Overall, eight soil health indicator's spatial CV ranged from 10 to less than 20%, and six indicator's spatial CV (excluding clay and silt) was less than 10% (Fig. 2−3 A). While soil texture components are typically not considered soil health indicators, clay and silt concentrations were included in the analysis for comparison with the soil health indicators: clay had similar spatial CV values as ACE, C:N, OM, and AWC, whereas silt had slightly higher spatial CV values than pH.

Additionally, Fine et al. (2017) demonstrated that soil texture properties impacted soil health indicator vales; many of the indicators exhibited dissimilar spatial and temporal CV values among coarse, medium, and fine soil textural categories (Supplementary Fig 2).

The temporal CV in the soil health indicators was calculated for control strips at 51–63 SHP locations that had complete soil health analysis data in at least two years with the same crop. The temporal CV for 13 of the 16 indicators was greater than 10% (Fig. 2−3 B). Solvita had the highest temporal CV and was 37 percentage points greater than OM and ACE. Soil pH temporal CV was approximately 27 percentage points lower than S in the chemical property category. Temporal CV in AWC and AggStabl had differences similar to those observed in the spatial CV. The temporal variation observed in clay and silt may be more related to sampling or analytic variability than actual changes in soil texture components over time. Specifically, the rapid determination soil texture methodology used on these soil samples is a modified version of the NRCS hydrometer methods and is known to have an analytical variability range of 0–6% CV (Kettler et al., 2001).

Soil subgroups influenced spatial CV values for clay concentration only (Fig. 2−3). In contrast, temporal CV was differentiated by suborder groups in WEON, C:N, S, Ca, and Mg. Across all the indicators, the mean temporal CV was 32% greater than mean spatial CV. The high P spatial CV and high S temporal CV likely resulted from elevated test values of these nutrients at select locations (Fig. 2−3). Hurisso et al. (2018) reported soil test analytic variability (CV) between 2.6 and 23% for ACE, microbial respiration (one-day incubation), and ActC, suggesting that analytic variability can account for some of the variation in the temporal and spatial CV values in this study. Corn and soybean yield had similar temporal and spatial CV (Fig. 2−4). Compared to the soil health indicators, the corn and soybean yield spatial CV was lower than all indicators except pH. In contrast, the yield temporal CV was similar in magnitude to half of the indicators. Among yield CV's, the soil suborders influenced only soybean temporal CV. Notably, the elevated temporal CV in suborder class 2 was observed in soybean yield as well as C:N, S, Ca, and Mg.

3.1.3. Correlations Among Soil Health Indicators

Correlation analysis was used to explore the relationships among soil health indicators temporal average soil test values (Table 2−3). Among 190 possible pairwise comparisons of 18 soil properties (indicators and soil texture), and two climate variables, 69% of those pairs were correlated ($p < 0.05$). Percent clay was correlated with 15 of 17 soil health indicators, OM was correlated with 13, while percent silt was correlated with only seven of 17 indicators. Moderately high correlations observed between OM and Ca ($r = 0.73$) and Mg ($r = 0.6$) but not with K may be explained by the common use of organic and inorganic fertilizers to manage K but not Ca and Mg, resulting in a possible decoupling of the correlations. Organic matter had positive correlations with ten soil health indicators and was most strongly correlated with ActC ($r = 0.84$), Ca ($r = 0.79$), and percent clay ($r = 0.69$). This evidence supports findings that soil OM is related to soil biological, chemical, and physical properties and functions (Krull et al., 2009; Murphy, 2014). Nunes et al. (2018) also reported that OM had high correlations with ActC, ACE, and Resp (4 d) ($r > 0.70$) in clay loam, silt loam, and loamy fine sand soils in the Northeastern U.S. Moderately strong correlations ($r > 0.60$) among OM, ACE, ActC, and

Resp (4 d) were reported by Fine, et al. (2017) in over 5700 soil samples across the Eastern and Midwestern US. Additionally, Franzluebbers & Pershing (2020) found high correlations ($r > 0.70$) among soil properties related to measures of microbiological activity in predominantly sandy loam soils; however, in the current study, ACE, Resp (4 d), and Solvita had low correlations (r = 0.19–0.31).

To determine how grouped soil texture classes impacted correlations among the indicators, further correlation analyses were conducted on the indicator's temporal means by coarse, medium, and fine texture categories (data not shown). While the strength of many correlations between indicators differed across the texture categories, there was no significant improvement in their correlations beyond what was observed in the aggregate dataset. The lack of improved strength in correlations when samples were grouped by similar soil texture, may be a consequence of variability among the indicator values and their associated soil textures across all SHP sites. Additionally, different sample sizes within each texture category make it difficult to assess the validity of how the correlation strength changes when the data is split by texture group because the variation in the data will be inherently different in a sample with 17 observations (coarse group) versus 169 (medium group) or 69 (fine texture group). These results emphasized that assessing the connection between condition and capability might be better situated at the site-specific level rather than across an entire region such as the Midwestern U.S.

In addition to the correlation analyses using the soil health indicator temporal average, the temporal CV allows further exploration of associations in temporal variation among soil indicators and climate variables (Supplemental Table 2−S1). These additional analyses are important for informing soil condition dynamics. Among 190 possible pairwise correlations, only 25% had a p-value < 0.05, with relatively low correlation coefficients (mean $r \approx |0.22|$), excluding the strongest correlations between the temporal CV of WEOC and WEON ($r = 0.70$), and GDD long-term average and GDD temporal CV ($r = 0.92$). Notably, WEOC and WEON were correlated with more soil health indicators and climate variables than the other indicators. Long-term climate variables (AWDR long-term average, temporal CV, GDD long-term average, and temporal CV) were positively or negatively correlated with the temporal CV values for six of 16 soil health indicators: P, Mg, AggStabl, ActC, Solvita, and WEON. However, the correlation values were low, indicating that long-term seasonal climate averages might not impact medium-term variation in some soil health indicators as much as other indicators. For example, temporal trends associated with changes in annual wetting and drying cycles and plant root and soil fauna activities influence soil bulk density and aggregate stability (Drewry et al., 2021) which may also be impacting soil biochemical acitivity and vice versa on a short-term scale than indicators like OM.

3.1.4. Multivariate Structure of Soil Health Indicator and Location Capability Factor Dataset

Principal component analysis (PCA) was utilized to better understand how soil health indicators and location capability factors account for variation in the dataset. Principal component analysis reduced the dataset's dimensions by simultaneously analyzing multiple variables to calculate correlations and summarize that co-variation into linear combinations called eigenvectors or principal components (PC) (Yeater et al., 2015). Unlike regression analyses which select variables that explain linear correlations of response and independent variables, PCA uses only independent variables to identify unobvious patterns in the dataset. This multivariate analysis provided insights into the relationships between variables that was not previously observable with bivariate analyses.

The PCA for soil health indicator temporal average values resulted in the first nine of 28 PC's having eigenvalues greater than 1, which indicated these first nine accounted for the total correlation among the dataset variables (Yeater et al., 2015). The first five PC's accounted for 60% of the dataset's total variation (Table 2−4). Dominant weights in PC1 related to C pools (e.g., OM, ActC), Ca, and percent clay, which corresponded to 24% of the dataset's variation. In PC2, climate variables (AWDR temporal CV, GDD temporal CV, GDD long-term average), AWC, and percent silt were dominantly weighted, corresponding to 14% of the dataset's variation. Field slope, K, WEOC, and WEON in PC3 were dominantly weighted, representing 9% of the variation. The tillage method and AWDR long-term average represented 8 and 6% of the dataset's variation in PC's 3 and 4, respectively. The AWDR long-term average was the most dominant in PC5 relative to other variables, accounting for 6% of the dataset's variation.

The PCA for soil health indicator temporal CV values had 10 PC's with eigenvalues greater than. Principal component 1 had similar weights on the variables as PC2 in the PCA on soil health indicator temporal average values (Table 2−4). Water-extractable organic N was the dominant variable in PC2, representing 10% of the variation in the soil health indicator temporal CV dataset. Principal components 3 and 4 accounted for 8% of the total variation (respectively). However, PC3 had dominant weights on four variables (site tillage, pH, P, and Resp (4 d)), while percent clay was the predominant variable in PC4. The relatively high weights on climate variables in PC1 emphasize the important influence climate has on variation in soil properties. The relatively large variation in PC2 indicates that WEON temporal variation is interconnected with other soil properties' temporal variation. These results have important implications for assessing soil biological health because WEON is an important energy source for soil microbial communities (Morrow et al., 2016).

In both PCA's, location climate, field conditions, and tillage methods frequently accounted for variation in the indicators, which demonstrated their importance when assessing soil health. Thus, when interpreting soil health test values, relationships among the location edaphic, climate, and management factors should be accounted for, as exemplified in the SMAF soil health score calculations and the Soil Health Assessment Protocol and Evaluation interpretive framework currently under development (Amsili et al., 2020; Andrews et al., 2004). Furthermore, the PCA results demonstrated that the climate variables have a relatively large impact on soil health indicators in aggregate that was not recognized by the separate bivariate correlations in the previous section (Table 2−3).

In a PCA conducted for soil health indicators in the CASH, Fine et al. (2017) suggested that cumulative variance shared by several PC's and the high dimensionality of the soil health indicator PC space (many PC's with eigenvalues >1) illustrated the complex nature of assessing soil health. Essentially, no single PC represented a significant proportion of their dataset's total variation so that each indicator importantly represents soil health, meaning many soil function indicators are needed to translate soil test values into information representing a living soil system (Doran & Parkin, 1994; Lal, 2016; Moebius-Clune et al., 2017). Furthermore, inclusion of environment factors in this

PCA demonstrated their strong ties to soil health assessment – a connection between soil capability and condition.

3.2. Sources of Cultivated Natural Capital Temporal Variation

Ordinary least squares best subsets multiple regression identified soil health indicators and location factors that helped explain the temporal trends and variation in cultivated natural capital, represented by corn and soybean yield across all SHP sites and years available for this analysis (Fig. 2−5 and 2−6; Supplementary Tables 2−S2 and 2−S3). Although these regression models were not intended to directly predict specific values of corn or soybean grain yield, the explanatory variables selected and their standardized coefficients in each model illustrated their relative explanatory power in the dependent variables. The most pronounced relative influence on yield variation was observed in soybean temporal CV due to the climate variables (Fig. 2−5). Field slope also had a large relative influence on soybean temporal CV. Additionally, ActC had a large relative influence on soybean temporal CV, indicating that ActC is an important soil health indicator related to temporal variation in soybean yield. In corn yield temporal average, K had a large relative standardized coefficient (Fig. 2−6), demonstrating the importance K has in photosynthate production (Havlin et al., 2014). This also supports a report of optimal soil test K ranging from 120–170 mg K kg^{-1} (Mallarino & Higashi, 2009), as the median value observed at SHP locations was 174 mg K kg^{-1}. Additionally, WEON was included in corn subsets only, further demonstrating the importance of this measure of organic N in corn production (Yost et al., 2018). It was also notable that ACE was not included in any of the eight models, nor was C:N included in the models with soil health indicators temporal CV values as explanatory variables. Fit and validation statistics illustrate models that used the temporal CV values for the dependent or soil health indicators explanatory variables had greater model error than those that used soil health indicators or yield temporal average values (Table 2−5).

Cumulatively, among all eight subsets explaining corn or soybean yield variation, growing season AWDR, field slope, moderate tillage, and four soil health parameters, Solvita, OM, Mg, and AWC were included in the subsets most frequently among their respective explanatory variable types (Fig. 2−7). These results demonstrate the important influence of climate, C cycling, and site field conditions on temporal variation in corn and soybean yield. For example, when estimating yield temporal average with soil health indicator temporal average values, microbial respiration (Solvita and Resp (4 d)), AWC, and C:N ratio were included for both crops. Differences among variables selected for inclusion in the models and the regression coefficients' standardized magnitude illustrate that yield temporal average and temporal CV are measures affected differently by climate, field, and soil factors.

3.3. Soil Connectivity and Codification: Implications of Soil Health Variation

3.3.1. Implications of Variation for Soil Health Indicator Interpretation and Connectivity

The differences in temporal and spatial CV among the soil health indicators had significant implications for soil health sampling intensity. For example, indicators with higher temporal and spatial CV (microbial respiration, WEOC and WEON) may need greater sampling intensity relative to indicators with lower variation (OM, C:N, and ACE). Furthermore, due to the lack of databases containing this information, common soil health assessments currently do not include descriptions of soil health indicator temporal or spatial variation for comparative purposes when reports are delivered to growers. Including estimates of variation will help practitioners know the range of possible inherent variation when assessing changes in soil health following management adaptations.

The temporal and spatial CV results, in conjunction with the correlation results, provided evidence of dynamic and complex interactions. Biological soil health indicators with low temporal and spatial CV (e.g., OM, C:N, ActC, ACE) had low correlations with indicators that had high temporal and spatial CV (Solvita, AggStabl). In contrast, OM and ActC had moderately high correlations with S, Ca, and Mg while having comparable temporal and spatial CV. However, C:N had low to weak correlations with OM, ActC, and ACE while having similar temporal and spatial CV. Undoubtedly, further insight into the drivers of variation and correlations among the soil health indicators, soil texture, and climate variables, from a complex systems perspective (F. C. Nunes et al., 2020; Yeater et al., 2015), might allow researchers to more effectively model and account for these relationships when recommending sampling intensity or when evaluating changes in assessment results.

The two PCA's revealed patterns that support the principles of soil formation (i.e., soil-forming factors are parent material, topographical relief, climate, vegetation soil biology and human impacts, and time): soil texture is a mediating property related to soil C pools, mineral element supply, and soil water availability; climate and weather variation are fundamental factors related to soil property variation; human crop and soil management due to crop rotations and nutrient amendments corresponds to temporal variation in biological and chemical soil properties (Table 2−4). These multidimensional patterns reinforce the concept that soil health assessment aims to characterize the complex soil ecosystem. The PCA results also support the preceding discussion of bivariate correlations and evaluation of the soil health indicators temporal CV. Generally, the PCA's provided evidence that the temporal average of many soil health indicators is related, yet their temporal CV is often different. These results demonstrated that some indicators vary on different time scales (F. C. Nunes et al., 2020) which adds complexity to soil health assessment interpretation when, for example, samples for all soil health indicators are collected and assessed on a composite basis simultaneously in the SMAF, CASH, and HSHT (Andrews et al., 2004; Haney et al., 2018; Moebius-Clune et al., 2017). Thus, further investigation is required to determine if accounting for those differences in temporal variation may bring clarity to soil health assessment interpretation.

Furthermore, the relationships between soil health indicator variation and yield identified in the best subsets regression models also had implications for soil connectivity and soil health assessment. Few of these indicators had strong relative effects on corn or soybean yield suggesting that a composite assessment of soil health may be more relatable to crop yield over time than individual relationships between indicators and yield. This finding was substantiated by four indicators (ActC, AWC, Ca, and K) that were selected in the corn yield temporal average regression model. These indicators also had dominant weights in the PCA on indicator temporal average values (Fig. 2−6, Table 2−4). Furthermore, those four indicators' temporal CV values were within a similar range of corn yield temporal CV (Fig. 2−3 and 2−4). Three of those indicators (excluding Ca) were also selected in the corn yield temporal CV models; however, none of them had dominant weights in the PCA on soil health indicator temporal CV values (Table 2−4;

Fig. 2−6). In sum, these results point out that biological (ActC), chemical (Ca, K), and physical (AWC) indicators are needed in soil health assessments to make a connection between soil capability and cultivated natural capital. Thus, when practitioners, researchers and land managers alike, make plans to assess soil condition, a suite of indicators is needed to represent soil capability.

Collective results in this study suggested that although temporal CV in soil health indicators may explain corn and soybean yield variation, it does not always translate to a correlation with the variation in the temporal CV of other indicators (Table 2−4, Supplemental Table 2−S1). Indeed, these results offer evidence of the difficulties facing soil health assessment and further support the call for research into the nature of complex interactions represented by soil health. Ultimately, soil health assessments seek to increase connectivity between land managers and their soil resources, however, these results imply that soil health indicator temporal and spatial variation from different soil taxonomic categories and textural properties may impact how practitioners differentiate inherent variability in soil capability versus actual soil condition changes.

3.3.2. Implications of Variation for Soil Codification in the Midwestern U.S.

Soil codification can take the form of initiatives like the SHP or NCRS cost share programs that lower entry cost to adopt soil health promoting practices, or public or private financial instruments that protect growers from crop failure. These analyses from soil health assessments linking soil condition variation to cultivated capital demonstrated that for corn and soybean yield temporal CV, the weather variation has a stronger effect than variation in soil properties (Fig. 2−7). An implication, especially for soybean production, is that growers may need to have greater awareness that soybean yields will become increasingly temporally variable as climate variation becomes more extreme. Although the effect from soil property temporal variations on crop yield is less than climate, institutions that underwrite crop failure insurance policies may also need greater awareness brought by soil health assessments that signal variation in soil condition from a changing climate. Additionally, soil security can be strengthened as codified knowledge (i.e., research or local knowledge of best practices) is exchanged between practitioner and grower and grower-to-grower. For example, increasing OM additions to the soil can increase yield and soil resilience to weather variation (Song et al., 2015). Furthermore, as soil health indicator datasets and models become more robust, crop yield variation attributed to year-over-year and within-season soil and climate variation can help practitioners become more resilient to weather by anticipating how non-normal weather might impact yield as demonstrated by Almaraz et al. (2008) or soil condition.

4. Conclusion

Although several studies have evaluated the capacity of soil health indicators to account for management differences (Cardoso et al., 2013; Hurisso et al., 2018; Morrow et al., 2016; Roper et al., 2017; van Es & Karlen, 2019), the objective of the current study was to examine field-scale soil health conditions across a wide geographic area. These analyses of soil health indicator short- to medium-term (1–5 years) data facilitate improved soil health monitoring by demonstrating typical values of soil health indicators, their spatial and temporal variation, and the relationship of that variation to crop yield. These descriptions illustrate the complexity of soil health assessment when soil health indicators vary differently in space and time and do not equally relate to variation in crop yield. These results can best support practitioners’ on-farm management by increasing their connectivity to the soil's condition and capability. Meaning that when there is greater recognition of inherent variation in soil health indicators, a grower's interpretation of a soil condition assessment may now lead to a greater willingness to persist with soil health improving practices when challenges arise from implementing the new adaptations. Furthermore, the descriptions provided an illustration that the range of variation in soil health indicators is higher for many of the biological than for chemical and physical properties. As well, the range of variation is dependent on soil taxonomy and texture. Future studies have an opportunity to investigate intervals in space and time for soil health sampling recommendations. Additionally, long-term (>10 years) monitoring is needed to establish definitive soil health temporal variation patterns.

Acknowledgments

The authors would like to express appreciation to the Soil health Partnership staff and field management team, whose work made this research possible. We also extend our deep appreciation to two anonymous reviewers and the editor-in-chief, whose comments provided guidance which greatly improved the manuscript's quality. Finally, we want to thank Briana Bowen of the Center for Anticipatory Intelligence, Utah State University, for her keen insight on policy implications.

5. Tables and Figures

Table 2−1

Soil health indicator abbreviations, units, description, analysis method, and method citation.

Soil health indicator	Units	Description	Soil function[a]	Soil security dimension[b]	Analysis method	Citation
Biological properties						
Organic matter (OM)	g kg^{-1}	Carbon based materials originating from living organisms	iv	Cn	Calculated as weight lost from a soil sample on ignition.	Schindelbeck et al., 2016
Permanganate oxidizable carbon (active carbon) (ActC)	mg kg^{-1}	A measure of easily available organic carbon energy source for soil microbes.	iii, iv	Cn	Photospectrometry analysis of oxidized potassium permanganate extractant.	Schindelbeck et al., 2016
Autoclaved citrate extractable soil protein index (ACE)	mg g^{-1}	A measure of organically bound nitrogen. Microbial activity makes this organic matter fraction available for plant use.	iii	Cn	High pressure and temperature extraction of citrate solution.	Schindelbeck et al., 2016
Soil microbial respiration 4-day incubation (Resp (4 d))	mg CO_2 C g^{-1}	A measure of soil microbial metabolic activity.	iv	Cn	Quantification of CO_2 gas trapped in solution evolved from re-wetted soil incubated 96 hours.	Schindelbeck et al., 2016
Soil microbial respiration 1-day incubation (Solvita)	mg CO_2 C kg^{-1}	A short duration measure of soil microbial metabolic activity.	iv	Cn	Paper chromatography quantification of CO_2 gas evolved from re-wetted soil incubated 24 hours.	(Haney, 2020; Ward Laboratories, 2020)
Water-extractable organic carbon (WEOC)	mg kg^{-1}	Measure of easily available organic carbon energy source for soil microbes.	iii, iv	Cn	Quantification of organic C extracted with water from a soil sample.	(Haney, 2020; Ward Laboratories, 2020)
Water-extractable organic nitrogen (WEON)	mg kg^{-1}	Measure of organically bound nitrogen. Considered as a “nutritional” source for microbes.	iii, iv	Cn	Quantification of organic N extracted with water from a soil sample.	(Haney, 2020; Ward Laboratories, 2020)
WEOC:WEON ratio (C:N)	—	Balance between energy and “nutrition” for soil microbes.	iii, iv	Cn	Ratio of WEOC to WEON.	(Haney, 2020; Ward Laboratories, 2020)
Chemical properties						
pH	—	Affects availability of nutrients and biological properties in the soil.	i, ii, iii	Cb, Cn	Voltage meter calibrated to determine Hydrogen ion activity in soil solution.	Watson and Brown, 1998
Soil chemical nutrients: P, K, S, Ca, Mg	mg kg^{-1}	Soil nutrients needed for healthy plant growth.	i, ii	Cn	Mehlich-III extractant method and quantified using inductively coupled atomic plasma spectroscopy.	Soil and Plant Analysis Council, 1999; Warncke & Brown, 1998
Physical properties						
Available water capacity (AWC)	g g^{-1}	A measure of soil water available for plant uptake.	i, ii	Cb	Amount of water extracted from a pulverized and sieved soil sample using a pressure chamber.	Schindelbeck et al., 2016
Soil aggregate stability (AggStabl)	%	Proportion of soil aggregates resistant to degradation following rain.	ii	Cn	Calculated from soil remaining on a 0.25 mm sieve following simulated rainfall.	Schindelbeck et al., 2016
Silt and clay	%	Soil proportions of particle size 0.002—0.05 mm (silt), and less than 0.002 mm (clay).	i–iv	Cb	Rapid 4-hour quantification of sand, silt, and clay from soil/water solution.	Schindelbeck et al., 2016

[a] i, biomass production; ii, storing, filtering, and transforming water, nutrients, substances; iii, provisioning for habitat; iv, carbon pool. [b] Cb, capability; Cn, condition.

Table 2−2

Units and descriptions for soil test, environment, and management variables utilized on 96 Soil Health Partnership locations.

Variable	Units or label	Description
AWDR LT avg	—	Abundant and well-distributed rainfall (AWDR) long-term average
AWDR TCV	%	AWDR long-term temporal coefficient of variation
GDD LT avg	$^{\circ}C$	Growing degree day (GDD) long-term average
GDD TCV	%	GDD long-term coefficient of variation
Corn/soybean	Mg ha^{-1}	Crop yield
Field tillage	0	No-till
	1	Strip tillage
	2	Vertical tillage
	3	Conventional disk/harrow tillage
Field slope	1	0-2%
	2	2-5%
	3	5-9%
Soil suborder classification	2	Suborders with moderate to high C sequestration potential (Alboll, Aquoll, Udoll, and Ustoll)
	3	Suborders with moderate to low C sequestration potential (Aqualf, Udalf, Ustalf)

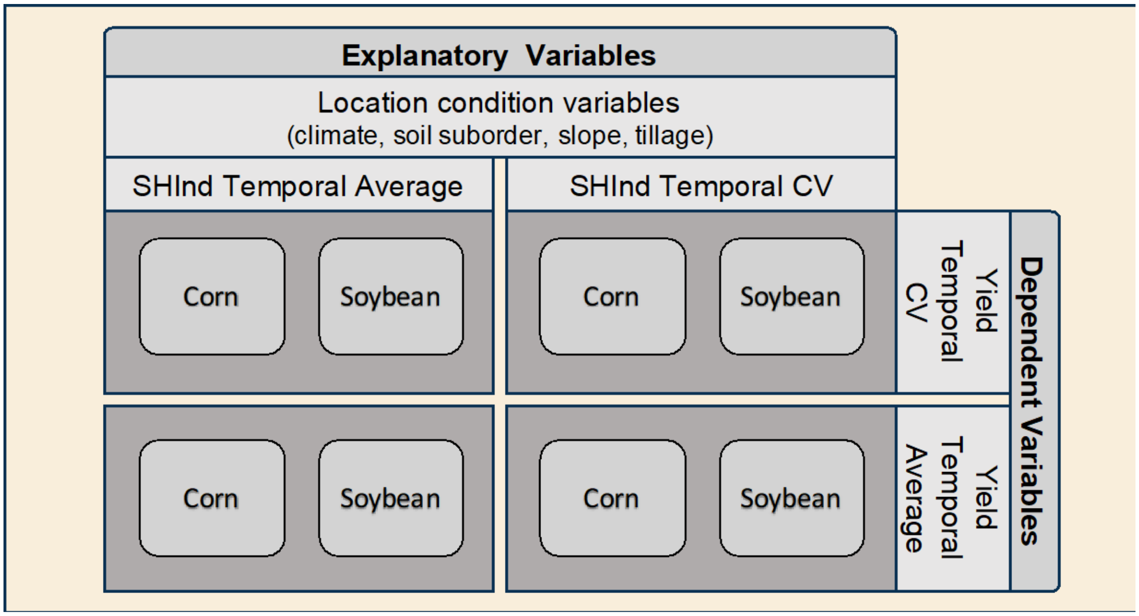

Fig. 2−1. Four combinations of explanatory and dependent variables used in the multiple regression analyses resulted in eight total models, four each for corn and soybean yield. CV, coefficient of variation; SHInd, soil health indicators.

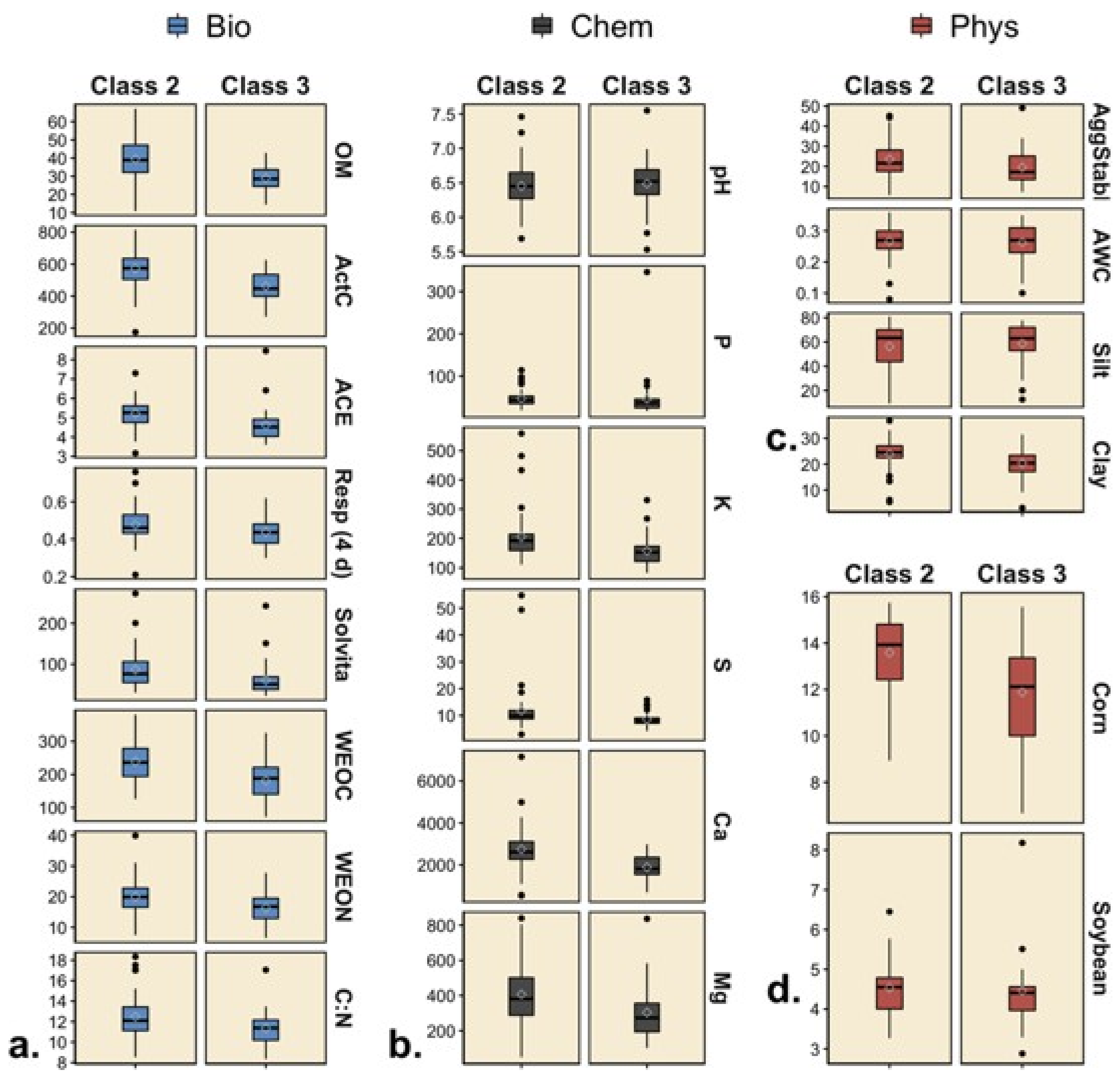

Fig. 2−2. Range of values for biological (a), chemical (b), and physical (c) soil properties, used as soil health indicators, and (d) yield (Mg ha^{-1}) at 94–96 Soil Health Partnership (SHP) sites. Soils were classified by their relative inherent potential for C sequestration (2, higher vs 3, lower) defined in the Soil Management Assessment Framework. The diamond shape represents the mean, the box represents the first quartile, median, and third quartile. See Table 2−1 for abbreviation descriptions and units. See Table 2−2 for soil class definitions.

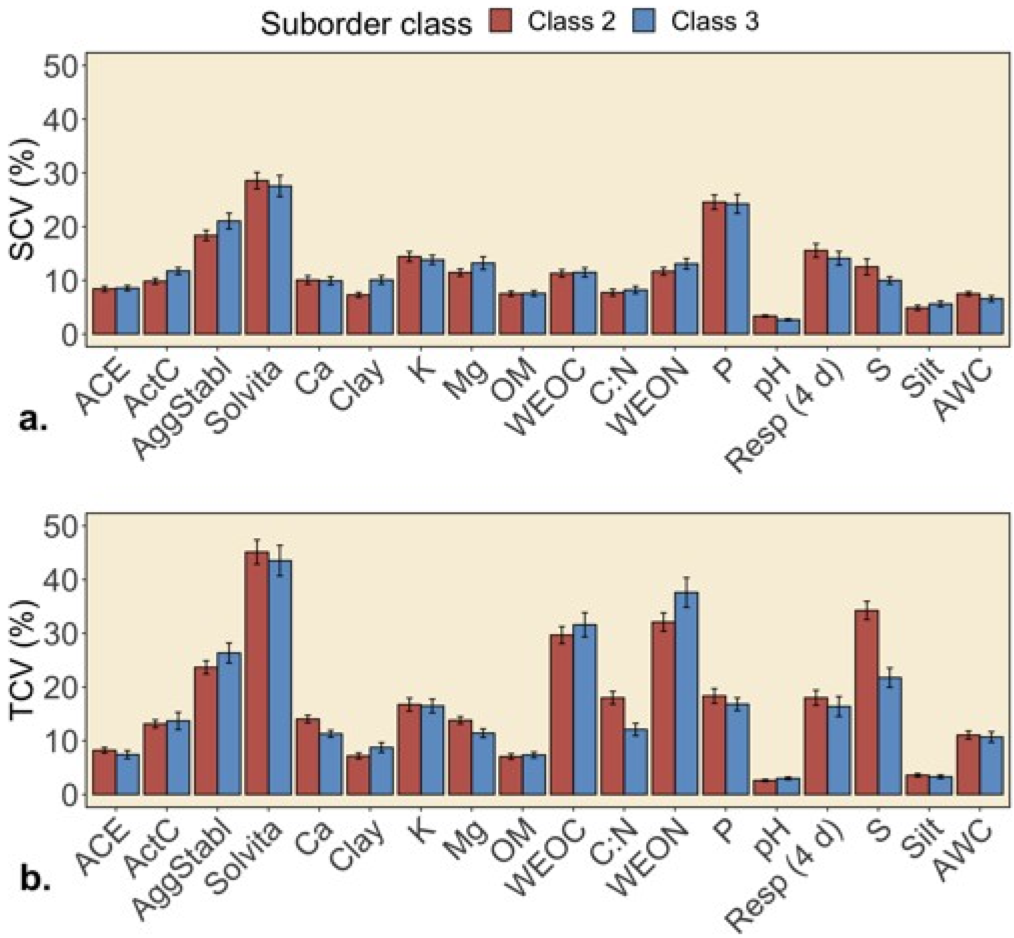

Fig. 2−3. Soil property (a) spatial coefficient of variation (CV), and (b) temporal CV. See Table 2−1 for abbreviation descriptions. Error bars indicate standard error of the mean.

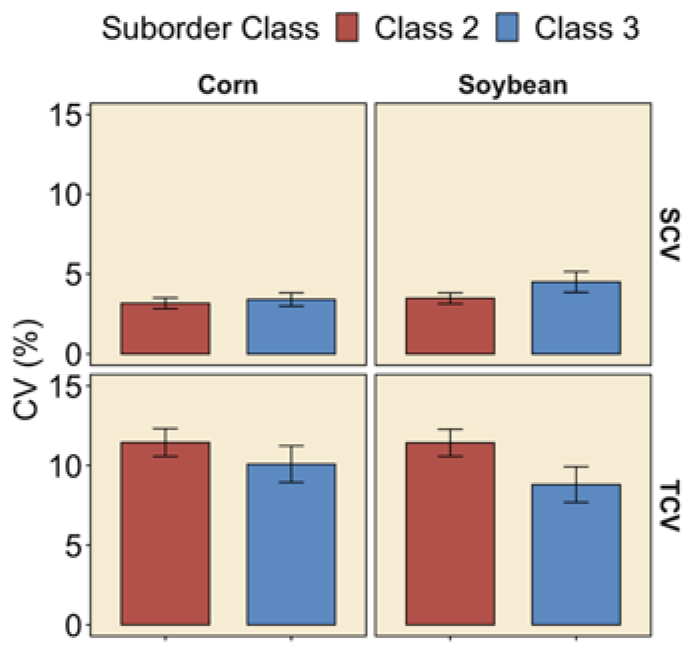

Fig. 2−4 Corn and soybean yield spatial and temporal coefficient of variation (SCV, TCV) for two soil suborder classes. See Table 2−2 for soil suborder class definitions. Error bars indicate standard error of the mean.

Table 2−3

Pearson correlation coefficients ($p < 0.05$) for soil health indicator temporal average values from 254 on-farm plots at Soil Health Partnership locations across the Midwestern US. See Table 2−1 for abbreviation descriptions.

	Silt	Clay	OM	pH	P	K	S	Ca	Mg	Agg	AWC	ActC	ACE	Resp	Solvita	WEOC	WEON	C:N	AWDR
Clay	0.38																		
OM	ns[a]	0.69																	
pH	-0.23	ns	ns																
P	-0.30	-0.30	ns	0.19															
K	ns	ns	0.24	ns	0.26														
S	ns	0.29	0.57	ns	0.28	0.30													
Ca	ns	0.73	0.79	0.19	ns	ns	0.40												
Mg	ns	0.60	0.60	0.20	ns	ns	0.32	0.48											
Agg	-0.39	0.31	0.47	ns	ns	ns	0.32	0.43	0.43										
AWC	0.80	0.42	0.17	-0.23	-0.36	0.17	ns	0.15	ns	-0.34									
ActC	0.14	0.58	0.84	ns	ns	0.29	0.53	0.67	0.52	0.36	0.24								
ACE	ns	0.18	0.58	-0.16	0.27	0.47	0.50	0.36	0.27	0.40	ns	0.61							
Resp	ns	0.36	0.38	0.21	-0.14	0.13	ns	0.43	0.35	0.23	ns	0.36	0.31						
Solvita	ns	0.32	0.42	ns	ns	0.16	0.19	0.40	0.14	0.13	ns	0.33	0.22	0.19					
WEOC	ns	0.44	0.55	ns	0.19	0.25	0.31	0.54	0.32	0.27	ns	0.44	0.33	0.24	0.42				
WEON	-0.16	0.26	0.37	ns	0.33	0.19	0.18	0.37	0.21	0.22	-0.22	0.29	0.27	0.13	0.42	0.87			
C:N	ns	0.31	0.34	ns	-0.13	0.14	0.38	0.31	0.24	ns	0.21	0.31	0.14	0.17	ns	0.21	-0.24		
AWDR[b]	ns	ns	-0.19	ns	ns	-0.42	ns	-0.14	ns	ns	-0.12	-0.18	-0.33	ns	-0.16	ns	ns	ns	
GDD[b]	0.46	ns	-0.17	-0.21	ns	ns	-0.23	ns	-0.47	-0.36	0.32	ns	-0.18	-0.17	ns	ns	-0.12	ns	ns

[a] ns, not significant.

[b] Long-term average.

Table 2−4
Principal component analysis (PCA) eigenvector weights and eigenvalues, the proportion of variation and cumulative variation of the first principal components (PC) for soil health indicator (SHInd) temporal average and temporal coefficient of variation (TCV) values along with location condition variables from 175 plots across the Soil Health Partnership. A PC weight of 0.3 (absolute value) was used as a baseline threshold to represent dominant variables (bolded values) within a PC. See Tables 2−1 and 2−2 for variable abbreviations and descriptions.

	PCA 1: SHInd temporal average					**PCA 2: SHInd temporal CV**			
Variable	**PC1**	**PC2**	**PC3**	**PC4**	**PC5**	**PC1**	**PC2**	**PC3**	**PC4**
OM	**0.357**	0.003	-0.014	0.012	0.034	0.113	-0.002	0.047	-0.136
Ca	**0.336**	0.001	-0.066	-0.193	0.059	-0.015	-0.281	0.141	0.294
ActC	**0.333**	0.033	-0.010	0.063	0.045	0.153	0.270	0.251	0.256
Clay	**0.305**	0.130	-0.137	-0.153	0.133	0.160	-0.193	0.169	**-0.471**
ACE	0.263	-0.025	0.115	0.264	-0.172	0.121	0.268	0.223	**0.307**
Suborder class 2	0.261	0.065	-0.004	-0.122	-0.140	0.141	0.003	0.002	-0.245
Mg	0.258	-0.130	-0.160	0.055	0.249	0.141	-0.289	0.230	0.217
S	0.252	-0.053	0.025	0.248	-0.011	-0.002	0.248	0.005	0.037
WEOC	0.251	-0.069	**0.318**	-0.160	0.017	-0.161	0.297	0.080	-0.065
AggStabl	0.197	-0.239	-0.025	-0.063	0.039	0.001	-0.112	-0.223	0.281
Solvita	0.191	0.031	0.115	-0.145	-0.103	-0.037	0.072	-0.190	0.034
Resp (4 d)	0.184	-0.025	-0.130	-0.037	0.176	0.162	0.056	**0.310**	0.246
WEON	0.181	-0.131	**0.392**	-0.193	-0.013	-0.149	**0.387**	0.083	-0.052
C:N	0.150	0.067	-0.164	0.113	0.030	0.067	0.218	0.029	-0.088
K	0.128	0.038	0.165	**0.320**	**-0.329**	-0.026	-0.226	0.217	-0.043
AWDR TCV	0.123	**0.317**	-0.070	-0.103	-0.040	**0.361**	-0.115	0.052	-0.078
AWC	0.083	**0.353**	-0.201	0.172	-0.048	-0.032	-0.130	-0.049	0.272
Silt	0.054	**0.386**	-0.087	0.108	-0.014	**0.331**	-0.161	0.075	-0.085
GDD TCV	0.052	**-0.407**	-0.150	0.113	0.063	**-0.415**	-0.056	0.161	-0.106
Slope class 2	0.028	-0.217	**-0.359**	0.030	**-0.376**	-0.284	-0.083	0.170	0.046
Site tillage 1	0.009	-0.024	-0.235	**-0.412**	-0.101	0.017	0.186	**0.350**	-0.163
P	0.007	-0.180	**0.356**	0.210	0.096	-0.067	-0.140	**0.310**	-0.181
Site tillage 0	0.001	0.060	-0.006	**0.475**	0.248	0.090	-0.163	-0.048	0.200
pH	-0.019	-0.182	0.003	-0.148	**0.312**	0.022	0.117	**0.347**	0.146
Slope class 1	-0.025	0.188	**0.351**	-0.027	0.301	0.271	0.251	-0.251	-0.055
GDD LT avg	-0.027	**0.412**	0.158	-0.170	0.005	**0.437**	0.104	-0.105	0.033
Site tillage 2	-0.062	-0.094	0.219	-0.151	**-0.343**	-0.155	-0.034	-0.245	0.111
AWDR LT avg	-0.103	-0.026	-0.121	-0.059	**0.420**	-0.082	0.035	-0.021	0.099
Eigenvalue	6.67	3.88	2.39	2.14	1.68	3.88	2.89	2.29	2.10
Proportion	0.24	0.14	0.09	0.08	0.06	0.14	0.10	0.08	0.08
Cumulative	0.24	0.38	0.46	0.54	0.60	0.14	0.24	0.32	0.40

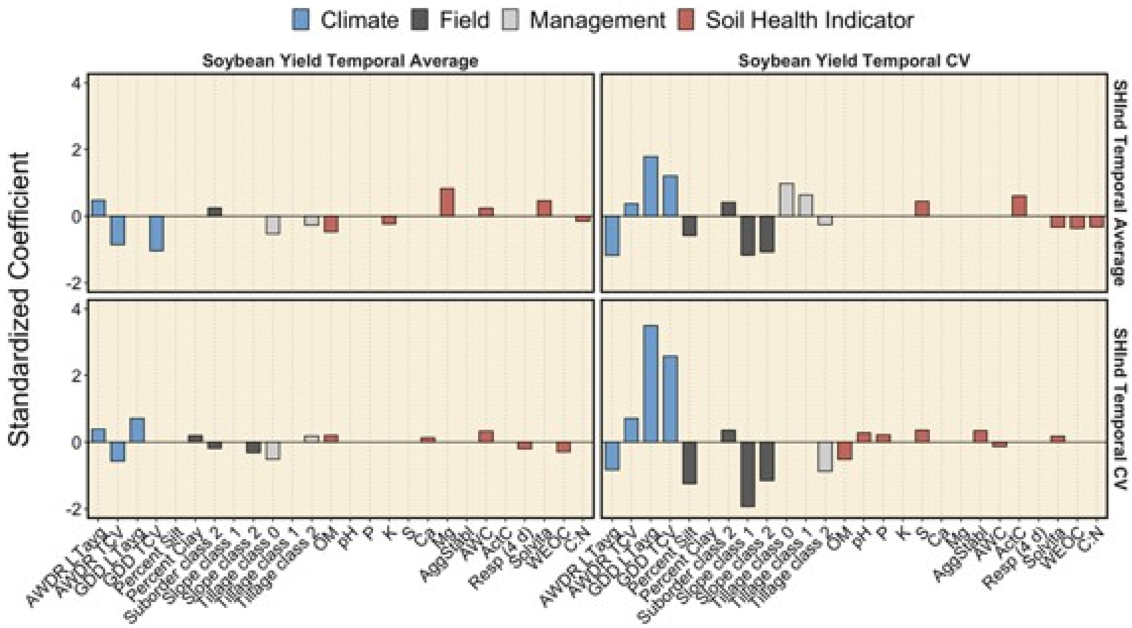

Fig. 2−5. Results of four regression models estimating soybean yield temporal average (left panels) or yield temporal coefficient of variation (CV) (right panels) utilizing location condition factors and soil health indicators (SHInd) as explanatory variables. Upper panels used the SHInd temporal average, and lower panels used the SHInd temporal CV values. Intercepts were set to 0 and were not shown. To represent the explanatory variables' relative impact on the dependent variable, the standardized coefficients estimate the number of standard deviations the dependent variable will change for one standard deviation change in an explanatory variable. See Table 2−1 and 2−2 for variable abbreviations and definitions.

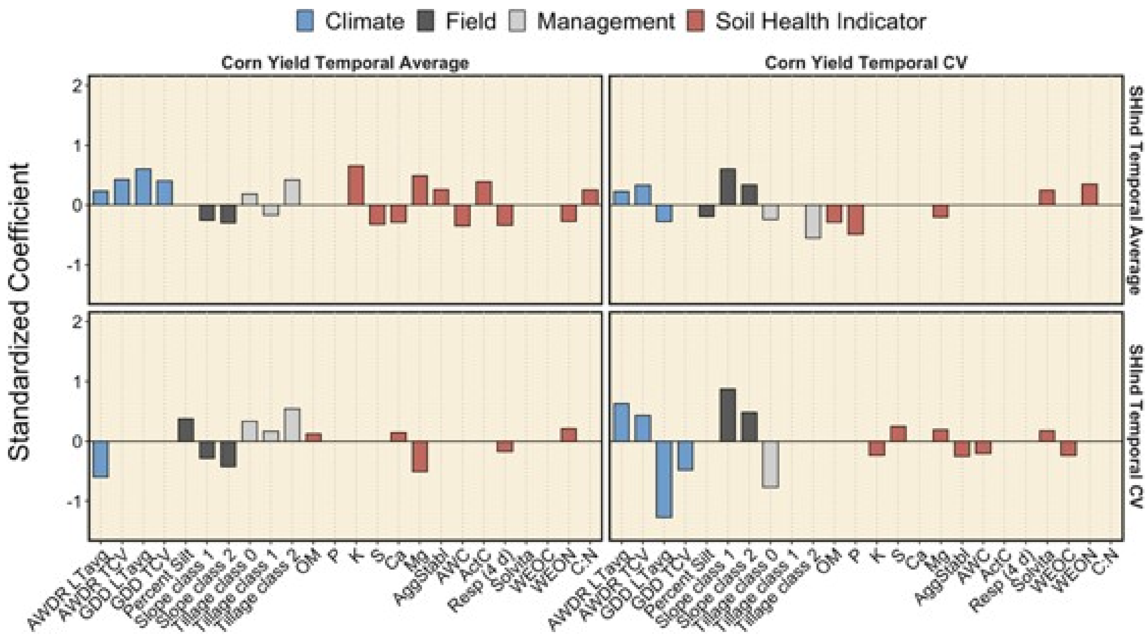

Fig. 2−6. Results of four regression models estimating corn yield temporal average (left panels) or yield temporal coefficient of variation (CV) (right panels) utilizing location condition factors and soil health indicators (SHInd) as explanatory variables. Upper panels used the SHInd temporal average, and lower panels used the SHInd temporal CV values. Intercepts were set to 0 and were not shown. To represent the explanatory variables' relative impact on the dependent variable, the standardized coefficients estimate the number of standard deviations the dependent variable will change for one standard deviation change in an explanatory variable. See Table 2−1 and 2−2 for variable abbreviations and definitions.

Table 2−5

Fit and validation statistics of eight multiple regression models estimating corn and soybean yield variation with the training (n = 175) or validation (n = 75) datasets.[a]

Model		Training set		Validation set			Obs. vs. Pred.
Dependent variable	Explanatory variables	$_{adj}R^2$	MSE	MSPE	MSPE:MSE	RMSE	$_{adj}R^2$
			Mg ha^{-1}	Mg ha^{-1}		Mg ha^{-1}	
Corn T avg	SHInd T avg	0.74	0.52	2.25	4.35	1.50	0.49
Corn T avg	SHInd TCV	0.67	0.62	2.66	4.32	1.63	0.42
Corn TCV	SHInd T avg	0.34	33.89	81.64	2.41	9.04	0.22
Corn TCV	SHInd TCV	0.55	21.47	60.73	2.83	7.79	0.51
Soy T avg	SHInd T avg	0.42	1.09	0.60	0.55	0.78	0.49
Soy T avg	SHInd TCV	0.84	0.08	0.35	4.17	0.59	0.52
Soy TCV	SHInd T avg	0.60	20.71	37.81	1.83	6.15	0.45
Soy TCV	SHInd TCV	0.78	13.19	63.44	4.81	7.97	0.43

[a] Abbreviations: adjR2, adjusted coefficient of determination; MSE, mean square error; MSPE, mean square prediction error; RMSE, root mean square error; Obs, observed; Pred, predicted; SHInd, soil health indicator; T avg, temporal average; TCV, temporal coefficient of variation.

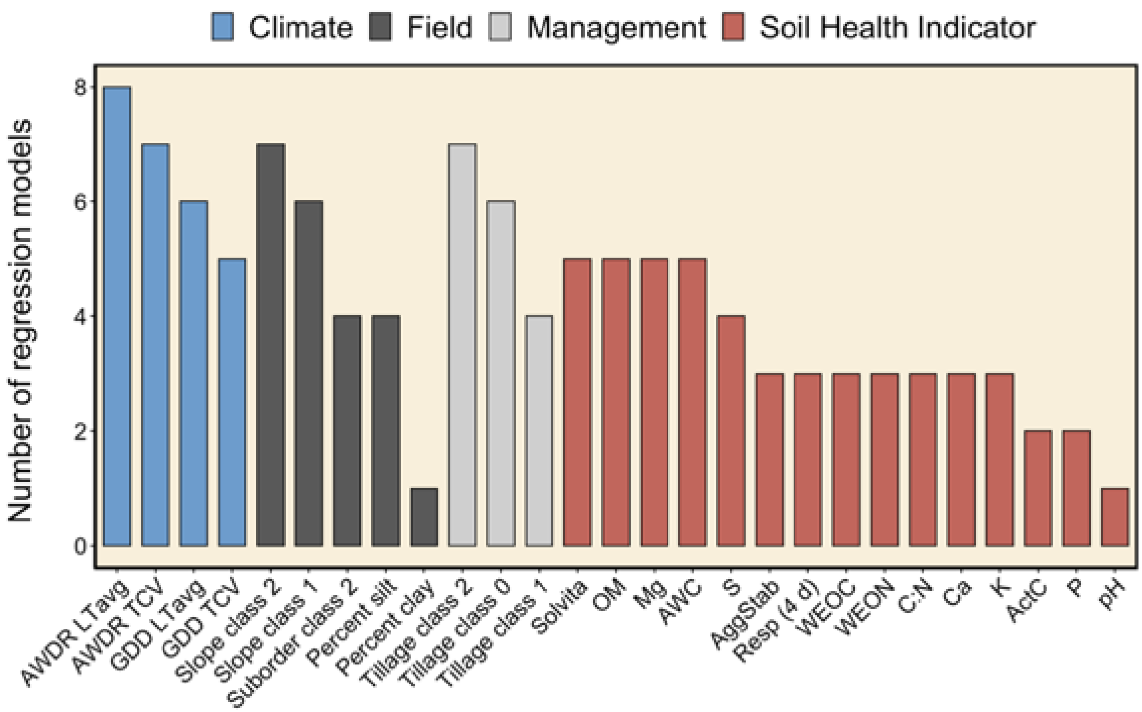

Fig. 2−7. Inclusion frequency of explanatory variables, grouped by type, selected in eight multiple regression models estimating corn and soybean yield variation (see Fig. 2−4 and 2−5). See Table 2−1 for variable definitions.

6. SUPPLEMENTARY MATERIAL

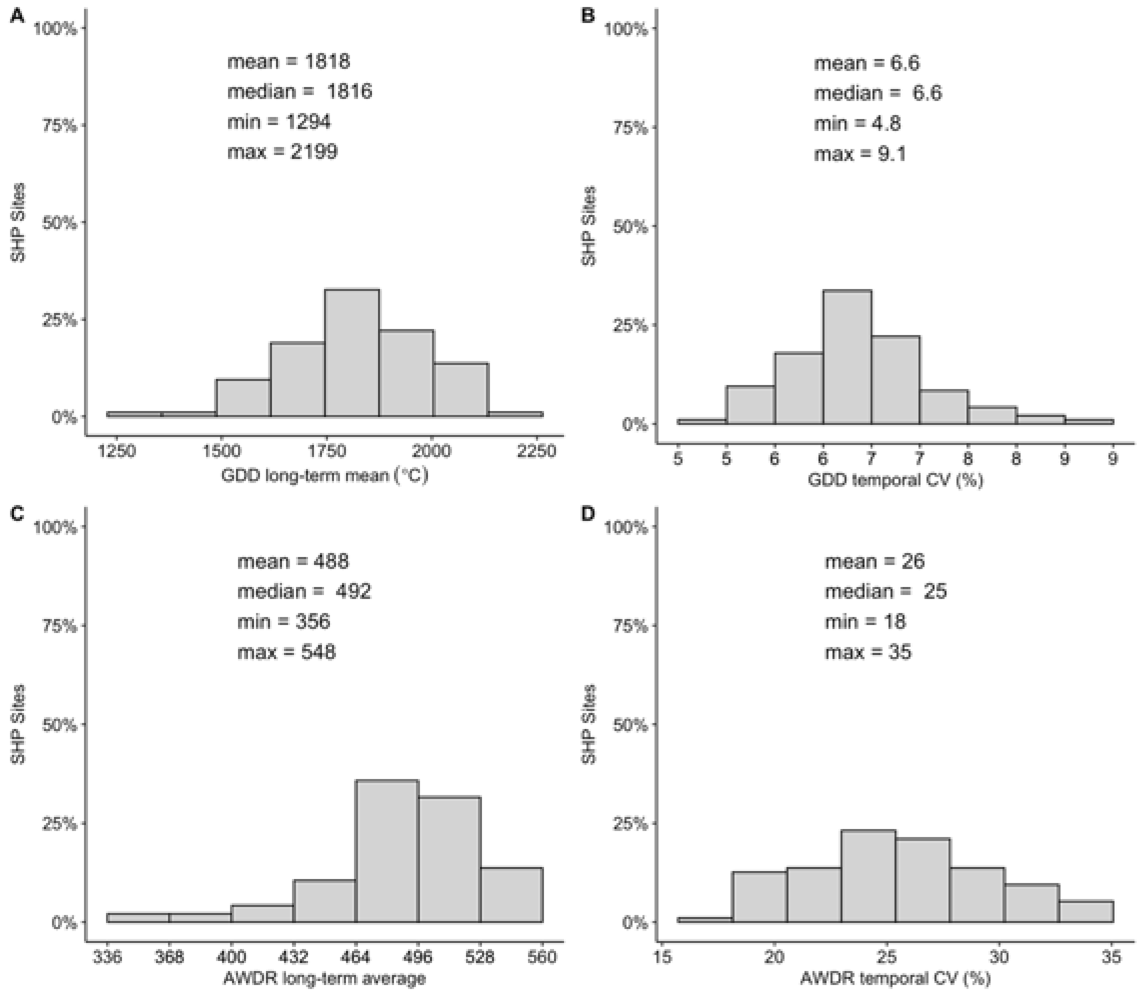

Fig. 2−S1. Distribution of seasonal weather variables (1 Apr – 31 Oct) for Soil Health Partnership sites. Growing degree day long-term mean (A), growing degree day temporal coefficient of variation (B), abundant and well-distributed rainfall (AWDR) long-term mean (c), and AWDR temporal coefficient of variation (D).

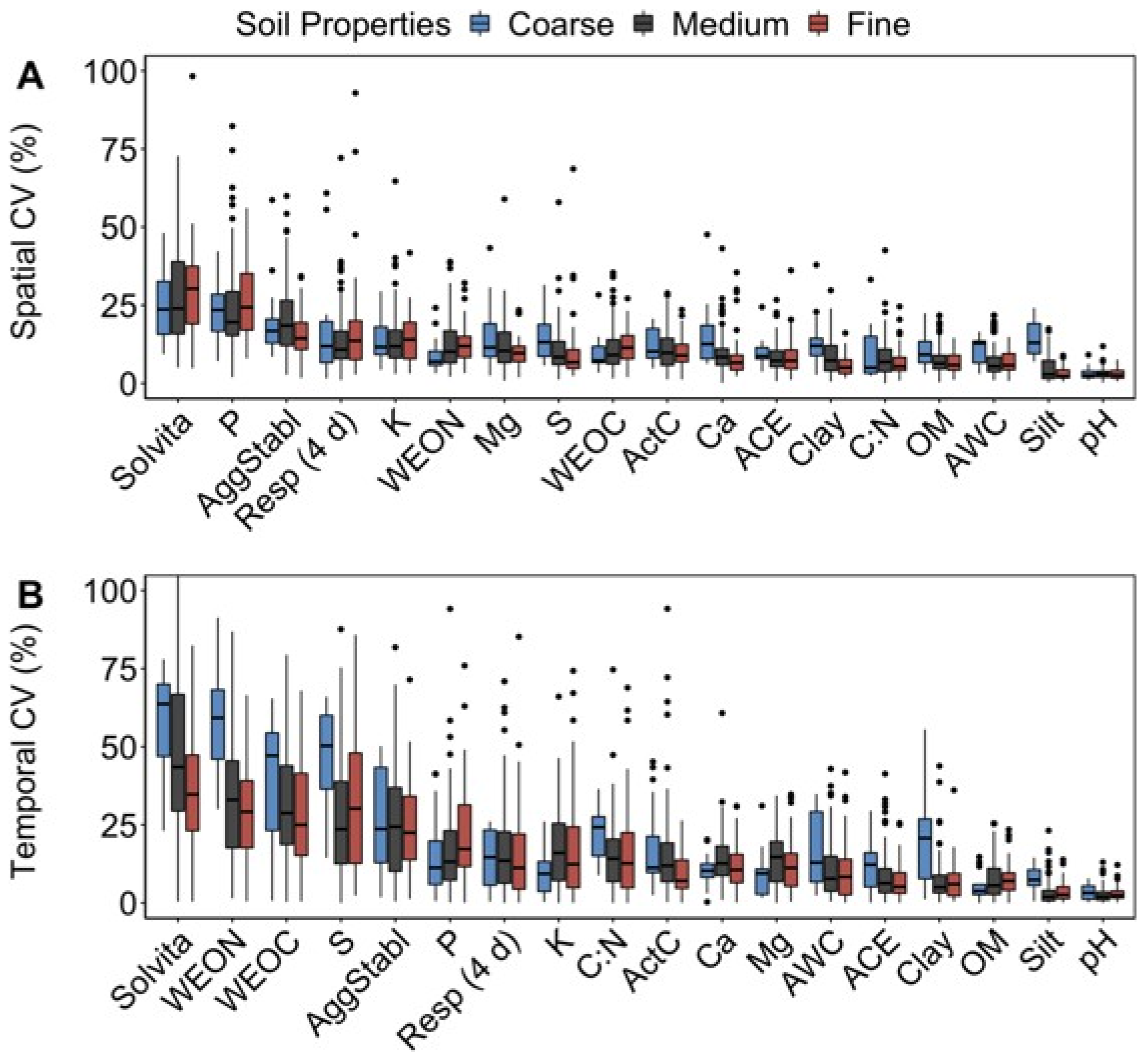

Fig. 2−S2. Spatial and temporal coefficient of variation for soil parameters at SHP sites. OM, organic matter; Agg, aggregate stability; AWC, available water content; ActC, active carbon; ACE, ACE soil protein; WEOC, water-extractable organic carbon; WEON, water-extractable organic nitrogen; C:N, carbon/nitrogen ratio. Soil texture groups were defined as, coarse: sand, loamy sand, sandy loam; medium: sandy clay loam, loam, silt loam, silt; fine: sandy clay loam, clay loam, silty clay loam, silty clay, clay.

Table 2‒S1

Pearson correlation coefficients ($p < 0.05$) for temporal coefficient of variation (CV) of soil health indicators in addition to climate variables.

	OM	pH	P	K	S	Ca	Mg	Agg	AWC	ActC	ACE	Resp	Solvita	WEOC	WEON	C:N	AWDR TCV	GDD TCV	AWDR LTavg
pH	0.15																		
P																			
K	0.22		0.46																
S																			
Ca		0.20		0.26	-0.14														
Mg				0.23		0.46													
Agg																			
AWC	-0.14																		
ActC		0.22																	
ACE		0.17								0.42									
Resp									0.16	0.17	0.20								
Solvita					0.16							-0.17							
WEOC					0.21	-0.24	-0.25	-0.16					0.17						
WEON						-0.20	-0.28	-0.14	-0.26	0.24				0.70					
C:N	-0.16												-0.14		0.27				
AWDR TCV										0.14					-0.22				
GDD TCV			0.19				-0.13	-0.16		-0.14					0.18		-0.50		
AWDR LTavg			-0.20				-0.25						-0.14						
GDD LTavg			-0.21					0.15		0.15							0.46	-0.92	

Table 2−S2

Corn yield best subsets multiple regression coefficients for location condition variables and soil health indicator (SHInd) datasets estimating yield temporal average (T avg) or yield temporal coefficient of variation (TCV).

Explanatory variable [a]	Explanatory variable dataset			
	SHInd T avg		SHInd TCV	
	Dependent variable		Dependent variable	
	Yield T avg	Yield TCV	Yield T avg	Yield TCV
	Coefficient			
Intercept	-11.77	3.39	20.52	89.47
AWDR TCV	0.16	0.63		0.91
GDD TCV	0.9			-5.27
AWDR LT avg	0.01	0.04	-0.02	0.1
GDD LT avg	0.01	-0.01		-0.07
Percent silt		-0.08	0.03	
Slope class 1	-0.72	8.75	-0.78	12.01
Slope class 2	-1.03	5.82	-1.35	7.76
Tillage class 0	0.54	-3.43	0.95	
Tillage class 1	-0.58		0.52	
Tillage class 2	1.59	-10.56	1.81	-12.92
OM		-0.22	0.03	
P		-0.06		
K	0.01			-0.11
S	-0.17			0.09
Ca	-0.001		0.03	
Mg	0.005	-0.01	-0.08	0.15
Agg	0.04			-0.1
AWC	-8.24			-0.15
ActC	0.005			
Resp (4 d)	-4.45		-0.02	
Solvita		0.05		0.05
WEOC				-0.09
WEON	-0.06	0.39	0.01	
C:N	0.14			

[a] AWDR LT avg, Abundant and well-distributed rainfall (AWDR) long-term average; AWDR TCV, AWDR long-term temporal coefficient of variation; GDD LT avg, Growing degree day (GDD) long-term average; GDD TCV, GDD long-term coefficient of variation, Slope class 1, 0–2 % slope; Slope class 2, 2–5 % slope; Tillage class 0, no-till; Tillage class 1, strip tillage; Tillage class 2, vertical tillage; OM, organic matter; Agg, aggregate stability; AWC, available water content; ActC, active carbon; ACE, ACE soil protein; WEOC, water-extractable organic carbon; WEON, water-extractable organic nitrogen; C:N, carbon/nitrogen ratio.

Table 2−S3

Soybean yield best subsets multiple regression coefficients for location condition variables and soil health indicator (SHInd) datasets estimating yield temporal average (T avg) or yield temporal coefficient of variation (TCV).

Explanatory variable [a]	Explanatory variable dataset			
	SHInd T avg		SHInd TCV	
	Dependent variable		Dependent variable	
	Yield T avg	Yield TCV	Yield T avg	Yield TCV
	Coefficient			
Intercept	12.5	-3.18	-75.89	-329.62
AWDR TCV	-0.26	-0.08	0.6	1.15
GDD TCV	-1.84		11.14	23.39
AWDR LT avg	0.02	0.01	-0.28	-0.19
GDD LT avg		0.003	0.08	0.16
Percent silt			-0.25	-0.55
Percent clay		0.03		
Suborder class 2	0.73	-0.28	6.55	5.88
Slope class 1			-17.6	-30.13
Slope class 2		-0.51	-17.88	-19.85
Tillage class 0	-1.44	-0.73	13.91	
Tillage class 1	-0.87		10.78	
Tillage class 2		0.35	-5.93	-18.02
OM	-0.07	0.03		-0.71
pH				0.95
P				0.17
K	-0.01			0.15
S			0.95	
Ca		0.02		
Mg	0.01			
Agg				0.15
AWC		0.02		-0.1
ActC	7.09		0.04	
Resp (4 d)		-0.01		
Solvita	0.01		-0.05	0.05
WEOC		-0.01	-0.04	
C:N	-0.11		-1.2	

[a] AWDR LT avg, Abundant and well-distributed rainfall (AWDR) long-term average; AWDR TCV, AWDR long-term temporal coefficient of variation; GDD LT avg, Growing degree day (GDD) long-term average; GDD TCV, GDD long-term coefficient of variation, Slope class 1, 0–2 % slope; Slope class 2, 2–5 % slope; Tillage class 0, no-till; Tillage class 1, strip tillage; Tillage class 2, vertical tillage; OM, organic matter; Agg, aggregate stability; AWC, available water content; ActC, active carbon; ACE, ACE soil protein; WEOC, water-extractable organic carbon; WEON, water-extractable organic nitrogen; C:N, carbon/nitrogen ratio.

7. REFERENCES

Almaraz, J. J., Mabood, F., Zhou, X., Gregorich, E. G., & Smith, D. L. (2008). Climate change, weather variability and corn yield at a higher latitude locale: Southwestern Quebec. *Climatic Change*, *88*(2), 187–197. https://doi.org/10.1007/s10584-008-9408-y

Amsili, J., Holan, S., Karlen, D., Moorman, T., Nunes, M., Parker, P., Seybold, C., van Es, H., Wills, S., & Veum, K. (2020). Soil Health Assessment Protocol and Evaluation (SHAPE). *Translating Visionary Science to Practice: ASA-CSSA-SSSA International Annual Meeting*.

Amundson, R., Berhe, A. A., Hopmans, J. W., Olson, C., Sztein, A. E., & Sparks, D. L. (2015). Soil and human security in the 21st century. *Science*, *348*(6235). https://doi.org/10.1126/science.1261071

Andrews, S. S., Karlen, D. L., & Cambardella, C. A. (2004). The Soil Management Assessment Framework. *Soil Science Society of America Journal*, *68*(6), 1945–1962. https://doi.org/10.2136/sssaj2004.1945

Baveye, P. C., Baveye, J., & Gowdy, J. (2016). Soil “ecosystem” services and natural capital: Critical appraisal of research on uncertain ground. *Frontiers in Environmental Science*, *4*(JUN), 1–49. https://doi.org/10.3389/fenvs.2016.00041

Cardoso, E. J. B. N., Vasconcellos, R. L. F., Bini, D., Miyauchi, M. Y. H., Santos, C. A. dos, Alves, P. R. L., Paula, A. M. de, Nakatani, A. S., Pereira, J. de M., & Nogueira, M. A. (2013). Soil health: looking for suitable indicators. What should be considered to assess the effects of use and management on soil health? *Scientia Agricola*, *70*(4), 274–289. https://doi.org/10.1590/S0103-90162013000400009

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), 1247–1250. https://doi.org/10.5194/gmd-7-1247-2014

de Mendiburu, F. (2020). *agricolae: Statistical Procedures for Agricultural Research Version 1.3-3*. https://cran.r-project.org/web/packages/agricolae/index.html

Doran, J. W., & Parkin, T. B. (1994). Defining and assessing soil quality. In J. W. Doran, D. C. Coleman, D. F. Bezdicek, & B. A. Stewart (Eds.), *Defining soil quality for a sustainable environment* (pp. 1–21). ASA, CSSA, and SSSA. https://dl.sciencesocieties.org/publications/books/abstracts/sssaspecialpubl/definingsoilqua/1

Doran, J. W., & Zeiss, M. R. (2000). Soil health and sustainability: managing the biotic component of soil quality. *Applied Soil Ecology*, *15*(1), 3–11. https://doi.org/10.1016/S0929-1393(00)00067-6

Drewry, J. J., McNeill, S. J., Carrick, S., Lynn, I. H., Eger, A., Payne, J., Rogers, G., & Thomas, S. M. (2021). Temporal trends in soil physical properties under cropping with intensive till and no-till management. *New Zealand Journal of Agricultural Research*, *64*, 223–244. https://doi.org/10.1080/00288233.2019.1684323

Fine, A. K., van Es, H. M., & Schindelbeck, R. R. (2017). Statistics, Scoring Functions, and Regional Analysis of a Comprehensive Soil Health Database. *Soil Sci. Soc. Am. J.*, *81*, 589–601. https://doi.org/10.2136/sssaj2016.09.0286

Franzluebbers, A. J., & Pershing, M. R. (2020). Soil‐test biological activity with the flush of CO 2: VIII. Soil type and management diversity. *Soil Science Society of America*

Journal, *84*, 1658–1674. https://doi.org/10.1002/saj2.20128

Friedrichsen, C. N., Hagen-Zakarison, S., Wulfhorst, J. D., & Friesen, M. L. (2021). Soil Health and Well-Being: Redefining soil health based upon a plurality of values. *Soil Security*, 100004. https://doi.org/10.1016/j.soisec.2021.100004

Haney, R. L. (2020). *Soil Health Tool (SHT) version 4.7*. https://www.nrcs.usda.gov/wps/PA_NRCSConsumption/download/?cid=nrcseprd338485&ext=pdf

Haney, R. L., Haney, E. B., Smith, D. R., Harmel, R. D., & White, M. J. (2018). The soil health tool—Theory and initial broad-scale application. *Applied Soil Ecology*, *125*, 162–168. https://doi.org/10.1016/j.apsoil.2017.07.035

Havlin, J. L., Tisdale, S. L., Nelson, W. L., & Beaton, J. D. (2014). *Soil Fertility and Fertilizers: An Introduction to Nutrient Management* (8th ed.). Pearson Education.

Hurisso, T. T., Culman, S. W., & Zhao, K. (2018). Repeatability and Spatiotemporal Variability of Emerging Soil Health Indicators Relative to Routine Soil Nutrient Tests. *Soil Science Society of America Journal*, *82*(4), 939–948. https://doi.org/10.2136/sssaj2018.03.0098

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2017). *Introduction to Statistical Learning with Applicaitons in R* (2nd ed.). Springer.

Jenny, H. (1994). *Factors of Soil Formation: A System of Quantitative Pedology*. Dover Publications, Inc.

Kettler, T. A., Doran, J. W., & Gilbert, T. L. (2001). Simplified Method for Soil Particle-Size Determination to Accompany Soil-Quality Analyses. *Soil Sci. Soc. Am. J.*, *65*, 849–852.

Krull, E. S., Skjemstad, J. O., & Baldock, J. a. (2009). Functions of Soil Organic Matter and the Effect on Soil Properties. *GRDC Final Report CSO00029*, 128.

Kutner, M., Nachtsheim, C., & Neter, J. (2004). *Applied Linear Regression Model* (4th ed.). McGraw-Hill/Irwin.

Kyveryga, P. M., Mueller, T. A., & Mueller, D. S. (2018). On-Farm Replicated Strip Trials. In D. K. Shannon, D. E. Clay, & N. R. Kitchen (Eds.), *Precision Agriculture Basics* (pp. 189–207). American Society of Agronomy, Crop Science Society of America, and Soil Science Socitey of America. https://doi.org/10.2134/precisionagbasics.2016.0096

Lal, R. (2016). Soil health and carbon management. *Food and Energy Security*, *5*, 212–222. https://doi.org/10.1002/fes3.96

Lawrence, M. A. B., & Friedlander, B. (2020, January 15). *Cornell, Nature Conservancy to study key climate projects*. Cornell Chronical. https://news.cornell.edu/stories/2020/01/cornell-nature-conservancy-study-key-climate-projects

Lehmann, J., Bossio, D. A., Knabner, I. K.-, & Rillig, M. C. (2020). The Concept and Future Prospects of Soil Health. *Nature Reviews Earth & Environment*, *1*, 544–553. https://doi.org/10.1038/s43017-020-0080-8

Mallarino, A. P., & Higashi, S. L. (2009). Assessment of Potassium Supply for Corn by Analysis of Plant Parts. *Soil Science Society of America Journal*, *73*(6), 2177–2183. https://doi.org/10.2136/sssaj2008.0370

McBratney, A. B., Field, D. J., & Koch, A. (2014). The dimensions of soil security. *Geoderma*, *213*, 203–213. https://doi.org/10.1016/j.geoderma.2013.08.013

McBratney, A. B., Field, D., Morgan, C. L. S., & Huang, J. (2019). On soil capability, capacity, and condition. *Sustainability*, *11*, 3350. https://doi.org/10.3390/su10023350

Mitchell, J. P., Shrestha, A., Mathesius, K., Scow, K. M., Southard, R. J., Haney, R. L., Schmidt, R., Munk, D. S., & Horwath, W. R. (2017). Cover cropping and no-tillage improve soil health in an arid irrigated cropping system in California's San Joaquin Valley, USA. *Soil and Tillage Research*, *165*, 325–335. https://doi.org/10.1016/j.still.2016.09.001

Moebius-Clune, B. N., D.J. Moebius-Clune, B.K. Gugino, O.J. Idowu, R.R. Schindelbeck, A.J. Ristow, H.M. van Es, J.E. Thies, H.A. Shayler, M.B. McBride, K.S.M Kurtz, D.W. Wolfe, & G.S. Abawi. (2017). *Comprehensive Assessment of Soil Health – The Cornell Framework* (3.2). Cornell University. http://soilhealth.cals.cornell.edu/training-manual/

Morrow, J. G., Huggins, D. R., Carpenter-Boggs, L. A., & Reganold, J. P. (2016). Evaluating Measures to Assess Soil Health in Long-Term Agroecosystem Trials. *Soil Science Society of America Journal*. https://doi.org/https://doi.org/10.2136/sssaj2015.08.0308

Murphy, B. W. (2014). *Soil Organic Matter and Soil Function – Review of the Literature and Underlying Data*. https://ecaf.org/wp-content/uploads/2021/02/Soil_Organic_Matter-Brian_Murphy.pdf

Norris, C. E., Bean, G. mac, Cappellazzi, S. B., Cope, M., Greub, K. L. H., Liptzin, D., Rieke, E. L., Tracy, P. W., Morgan, C. L. S., & Honeycutt, C. W. (2020). Introducing the North American project to evaluate soil health measurements. *Agronomy Journal*, *112*(4), 3195–3215. https://doi.org/10.1002/agj2.20234

North Dakota Agriculture Weather Network. (2019). *Corn Growing Degree Days (GDD)*. https://ndawn.ndsu.nodak.edu/help-corn-growing-degree-days.html

Nunes, F. C., Alves, L. de J., Carvalho, C. C. N., Gross, E., Soares, M., & Vara Prasad, M. N. (2020). Soil as a complex ecological system for meeting food and nutritional security. In *Climate Change and Soil Interactions* (pp. 229–269). Elsevier.

Nunes, M. R., van Es, H. M., Schindelbeck, R., Ristow, A. J., & Ryan, M. (2018). No-till and cropping system diversification improve soil health and crop yield. *Geoderma*, *328*(April), 30–43. https://doi.org/10.1016/j.geoderma.2018.04.031

Nunes, M. R., Veum, K. S., Parker, P. A., Holan, S. H., Karlen, D. L., Amsili, J. P., Es, H. M., Wills, S. A., Seybold, C. A., & Moorman, T. B. (2021). The soil health assessment protocol and evaluation applied to soil organic C. *Soil Sci. Soc. Am. J.*, *85*, 1196–1213. https://doi.org/10.1002/saj2.20244

Oakridge National Laboratory. (2020). *DAYMET Daily Surface Weather and Climatological Summaries*. Single Pixel Extraction Tool. https://daymet.ornl.gov/single-pixel/

Rinot, O., Levy, G. J., Steinberger, Y., Svoray, T., & Eshel, G. (2019). Soil health assessment: A critical review of current methodologies and a proposed new approach. *Science of the Total Environment*, *648*, 1484–1491. https://doi.org/10.1016/j.scitotenv.2018.08.259

Roper, W. R., Osmond, D. L., Heitman, J. L., Wagger, M. G., & Reberg-Horton, S. C. (2017). Soil Health Indicators Do Not Differentiate among Agronomic Management Systems in North Carolina Soils. *Soil Science Society of America*

Journal, *81*(4), 828–843. https://doi.org/10.2136/SSSAJ2016.12.0400

Schindelbeck, R. R., Moebius-Clune, B. N., Moebius-Clune, D. J., Kurtz, K. S., & van Es, H. M. (2016). *Cornell University Comprehensive Assessment of Soil Health Laboratory Standard Operating Procedures*.

Six, J., Frey, S. D., Thiet, R. K., & Batten, K. M. (2006). Bacterial and Fungal Contributions to Carbon Sequestration in Agroecosystems. *Soil Science Society of America Journal*, *70*(2), 555–569. https://doi.org/10.2136/sssaj2004.0347

Soil Survey Staff. (2019). *Web Soil Survey*. Natural Resources Conservation Service, United States Department of Agriculture. http://websoilsurvey.sc.egov.usda.gov/

Song, Z., Gao, H., Zhu, P., Peng, C., Deng, A., Zheng, C., Mannaf, M. A., Islam, M. N., & Zhang, W. (2015). Organic amendments increase corn yield by enhancing soil resilience to climate change. *Crop Journal*, *3*(2), 110–117. https://doi.org/10.1016/j.cj.2015.01.004

Stott, D. E., Andrews, S. S., Liebig, M. A., Wienhold, B. J., & Karlen, D. L. (2010). Evaluation of β-Glucosidase Activity as a Soil Quality Indicator for the Soil Management Assessment Framework. *Soil Science Society of America Journal*, *74*(1), 107–119. https://doi.org/10.2136/sssaj2009.0029

Svoray, T., Hassid, I., Atkinson, P. M., Moebius-Clune, B. N., & van Es, H. M. (2015). Mapping Soil Health over Large Agriculturally Important Areas. *Soil Science Society of America Journal*, *79*(5), 1420. https://doi.org/10.2136/sssaj2014.09.0371

Tremblay, N., Bouroubi, Y. M., Bélec, C., Mullen, R. W., Kitchen, N. R., Thomason, W. E., Ebelhar, S., Mengel, D. B., Raun, W. R., Francis, D. D., Vories, E. D., & Ortiz-Monasterio, I. (2012). Corn response to nitrogen is influenced by soil texture and weather. *Agronomy Journal*, *104*(6), 1658–1671. https://doi.org/10.2134/agronj2012.0184

Valani, G. P., Vezzani, F. M., & Cavalieri-Polizeli, K. M. V. (2020). Soil quality: Evaluation of on-farm assessments in relation to analytical index. *Soil and Tillage Research*, *198*(December 2019), 104565. https://doi.org/10.1016/j.still.2019.104565

van Es, H. M., & Karlen, D. L. (2019). Reanalysis Validates Soil Health Indicator Sensitivity and Correlation with Long-term Crop Yields. *Soil Science Society of America Journal*, *83*(3), 721–732. https://doi.org/10.2136/sssaj2018.09.0338

Ward Laboratories. (2020). *Haney Test Interpretation Guide v1.0*. https://www.wardlab.com/wp-content/uploads/2020/06/Haney-Rev-1.0-Interpretation-Guide-PDF.pdf

Warncke, D., & Brown, J. R. (1998). Potassium and Other Basic Cations. In Recommended Chemical Soil Test Procedures for the North Central Region Publication No. 221 (revised) (pp. 31–33). University of Missouri Ag. Exp. Station.

Wicklin, R. (2018). *Standardized regression coefficients*. The DO Loop - SAS Blogs. https://blogs.sas.com/content/iml/2018/08/22/standardized-regression-coefficients.html

Yeater, K. M., Duke, S. E., & Riedell, W. E. (2015). Multivariate Analysis: Greater Insights into Complex Systems. *Agronomy Journal*, *107*(2), 799–810. https://doi.org/10.2134/agronj14.0017

Yost, M. A., Veum, K. S., Kitchen, N. R., Sawyer, J. E., Camberato, J. J., Carter, P. R., Ferguson, R. B., Fernández, F. G., Franzen, D. W., Laboski, C. A., & Nafziger, E. D. (2018). Evaluation of the Haney Soil Health Tool for corn nitrogen recommendations across eight Midwest states. *Journal of Soil and Water Conservation*, *73*(5), 587–592. https://doi.org/10.2489/jswc.73.5.587

CHAPTER 3

RELATIONSHIPS AMONG SOIL HEALTH INTERPRETIVE FRAMEWORKS, SOIL TEXTURE, AND CROP YIELD[2]

Abstract

Soil health assessments are interpretive frameworks that estimate a soil's health by scoring various soil biological, chemical, and physical attributes to guide soil and crop management. Few large-scale analyses of soil health assessment scores exist. Thus, our objectives were to 1) summarize soil health scores at farms across much of the Midwestern U.S., 2) evaluate how individual soil health indicators influence assessment composite scores, 3) assess correlations among composite scores, and 4) determine the strength of significant correlations between soil health assessment scores and crop yield at three spatial and temporal scales, a) individual site-year, b) individual site, and c) all sites and years combined. Soil health and yield data were collected from 96 Soil Health Partnership farmers across nine Midwestern states over two to five years. Soil texture influenced soil health indicator values and scores in the study region. Correlation strengths among the Comprehensive Assessment of Soil Health (CASH), the Soil Management Assessment Framework (SMAF), and the Haney Soil Health Tool (HSHT) composite scores were soil texture dependent. Multiple years of data at individual sites improved the statistical correlations between yield and soil health scores compared to

[2] Coauthors: Matt A. Yost, Maria Bowman, Kristen Veum; A manuscript submitted to *Soil Science Society of America Journal*.

individual site years. When practitioners judge how well crop yield relates to soil health scores, examining data over time should enhance relationships. These results demonstrate that a multi-year commitment greatly improves soil health monitoring at the site level.

Core ideas:

- Soil texture effects on soil health scores were evaluated.
- Relationships between scores and yield were analyzed at three spatial and temporal scales.
- Composite scores are differentially influenced by individual indicators based on soil texture.
- Multiple years of data are needed to observe yield-score correlations at an individual site.
- Unique site-level factors influence yield and score correlations.

Abbreviations: ACE, autoclaved citrate extractable protein index; ActC, active carbon; AGG, wet aggregate stability; AWC, available water capacity; K, potassium; OM, organic matter loss on ignition; P, phosphorus; pH; Resp (24 hr), microbial respiration 24-hour incubation; Resp (96 hr), microbial respiration 96-hour incubation; SOC, soil organic carbon; WEOC, water-extractable organic carbon; WEON, water-extractable organic nitrogen.

1 | INTRODUCTION

As a concept of productivity and capability, soil health is “the capacity of soil to function as a vital living system to sustain biological productivity, maintain environmental quality and promote plant, animal, and human health” (Doran & Zeiss, 2000). This concept of healthy soil must be captured and conveyed through proper measurements sensitive to changes in soil processes and should represent connections among the soil, environment, and influences of human management (Andrews et al., 2004; Fine et al., 2017; Moebius-Clune, et al., 2016). Soil health interpretive frameworks, otherwise known as soil health assessments, were developed to estimate soil health from various soil properties, known as soil health indicators, to signal quantitative information to growers and practitioners about their soil's health status. These assessments translate an observed value from a property analyzed in a soil sample into a unitless score to interpret inherent potential for soil health and dynamic responses to management practices (Stott, 2019). Some scoring methods are based on site-specific conditions (climate and soil type) and crop factors (Andrews et al., 2004), while others are based on regional and textural categories (Fine et al., 2017), soil type and climate peer groups (M. R. Nunes et al., 2021), or soil property thresholds (Haney et al., 2018). Understanding relationships among the types of assessments and their relationship to soil health outcomes, such as crop yield, is crucial to utilizing and interpreting soil health assessments.

Traditional soil analyses report the concentrations of soil chemical properties for fertilizer recommendations. However, most soil health assessments characterize different combinations of soil biological, chemical, and physical properties and processes representing the soil as a living system (Karlen et al., 1997; Moebius-Clune et al., 2016). Originally, soil health measurements began as a home kit from the USDA-Natural Resources Conservation Service (Karlen et al., 1997). Now, suites of soil health analyses are available from several commercial labs across the U.S. The information from soil health assessment frameworks can be used to identify constraints to healthy soil, such as surface or subsurface hardness or low microbial activity, and then guide growers on management practices known to improve soil conditions (Moebius-Clune et al., 2017). However, a crucial need is to establish how inherent soil conditions such as sand, silt, and clay concentrations (i.e., soil texture) influence assessment composite score interpretation.

Several soil health interpretive frameworks utilize various soil health indicators to provide an overall soil health interpretation based on all the indicators measured, known as an index or composite score. Although the mathematical methods to translate a laboratory measurement into a unitless score differ, the frameworks are typically based on three general scoring curves: more-is-better (e.g., soil organic carbon), less-is-better (e.g., surface and subsurface penetration resistance), and midpoint-optimum (e.g., nutrients such as phosphorus where low or high concentrations may be detrimental to plant growth or environmental quality) (Andrews et al., 2004; Moebius-Clune et al., 2016). This general scoring approach is based on decades of soil research correlating soil properties with functions that lead to crop productivity and healthier soils (Andrews et al., 2004; Moebius-Clune et al., 2016; Stott, 2019).

One of the first interpretive frameworks was the Soil Management Assessment Framework (SMAF), developed in the early 2000s (Andrews et al., 2004). This framework interprets lab values using logic statements and algorithms parameterized by in-field characteristics such as soil texture, soil taxonomy, climate regime (annual average precipitation and temperature), topographical slope, crop type, and expert opinion (Andrews et al., 2004; Stott et al., 2010; Wienhold et al., 2009). Although researchers have shown that the SMAF effectively detects differences in soil management (Cherubin et al., 2016; Hammac et al., 2016), SMAF scores are not widely used by farmers. Low adoption by growers or commercial soil testing labs might be due to the lack of a user-friendly, publicly available interface or the need for additional details regarding field characteristics before translating soil measurements into soil health scores using the algorithms. In addition, SMAF was developed using a relatively small dataset and was not parameterized for a wide range of soils or climate regimes.

A second framework, Cornell University's commercialized Comprehensive Assessment for Soil Health (CASH), built upon concepts in the SMAF but utilized a statistical cumulative normal distribution, calculated from the mean and standard deviation of a calibration dataset, then converted into a percentile score (Moebius-Clune et al., 2016; Fine et al., 2017). Additionally, the CASH replaced some SMAF indicators and expanded the indicator suite to enable rapid testing in high throughput laboratory settings. In contrast to the SMAF that uses site-specific characteristics, the CASH assumes the calibration dataset comes from a random population of soil samples representing a wide range of soil characteristics within a region. Specific CASH indicators are also adjusted for soil textural class (Moebius-Clune et al., 2016; Fine et al., 2017). An advantage of the CASH is the simplicity of interpretation. However, a challenge to a broader application has been that a large proportion of the soil samples in the original calibration dataset came from Northeastern and Midwestern U.S. (Fine et al., 2017; Stott, 2019); however, the sample database is expanding.

A third common framework is the Haney Soil Health Tool or Haney Test (HSHT), based on soil biological and chemical measurements that indicate microbial activity, nutrient cycling, and potentially mineralizable nutrients. This framework's output has been used in nutrient recommendations for soil health improvement – a feature lacking in other frameworks (Haney et al., 2018; Harmel & Haney, 2013). The HSHT calculates an overall soil health score without scoring individual indicators as in the SMAF and CASH. A recent evaluation of the HSHT found that the 24-hour microbial respiration test had the strongest correlation ($R^2 = 0.8$) to the economically optimum nitrogen rate applied in corn (*Zea mays*, L.) (Yost et al., 2018). This result suggested that the microbial respiration test by itself may be a good predictor of appropriate N requirements versus the full HSHT assessment. Although the HSHT has become popular with growers and is increasingly available at soil analytical labs around the country, more research is needed to establish the relationship between HSHT score and differing management practices (Stott, 2019; Yost et al., 2018).

Each of the soil health indicators and assessment frameworks has the burden to demonstrate their relevance in diverse environments and for many cropping systems. Evaluation of agreement among scoring approaches is needed, given the differences in how they translate indicator lab values into a soil health interpretation (composite score). In addition, more research is required to quantify the relationships between soil health indicator assessment scores and crop yield. The objectives of this study were to 1) summarize soil health scores for a group of farms in the Midwestern U.S., 2) evaluate how individual soil health indicators influence assessment composite scores, 3) assess correlations among composite scores, and 4) determine the strength of correlations between soil health assessment scores and crop yield at three spatial and temporal scales, a) individual site-year, b) individual site, and c) all sites and years combined.

2 | MATERIALS AND METHODS

2.1 | Data Collection from On-Farm Evaluations of Soil Health

Soil Health Partnership on-farm trials began in 2014 with 14 locations in five states. At the closing of the program in 2021, there were over 200 sites in 16 states. The on-farm trials evaluating soil health-promoting practices included randomized strips of the growers' historical management as a control treatment and a soil health treatment that mainly included cover crops or reduced tillage, each replicated four times. All analyses in this study used data from the non-treated strips to examine the soil health assessment frameworks and scores without interference from treatments. Evaluation of soil health indicator temporal and spatial variation is available in Crookston et al. (2021), and an analysis of cover crop impacts on soil health indicators at SHP locations was reported by Wood and Bowman (2021). Further publication of treatment impacts on indicators and soil health assessment scores at SHP locations is forthcoming. The dataset used in this study included 96 SHP sites covering fields added to the network between 2014 and 2018 and their data from 2015 to 2019. This dataset was limited to sites that used annual rotations of corn and soybean (Glycine max). Grain yield data were collected annually from combine-mounted yield monitoring systems with corresponding global positioning system locations. Yield data quality assurance followed the Iowa Soybean Association's procedure (Kyveryga et al., 2018). Combine yield data were averaged for each strip on each farm.

Sampling for soil nutrient analyses at SHP sites occurred approximately annually between 2014 and 2019. Samples for soil health analyses were collected at SHP locations roughly every other year beginning in 2015. Samples for soil health indicators were collected from SHP sites in the spring before planting. Sampling points were predetermined and geolocated in a 0.4 ha grid across strips. Strips ranged in size from

0.4–4 ha. Samples for soil P, K, organic matter (OM), and pH were collected to the 0–5 cm and 5–15 cm depths. In addition, samples for soil active carbon (ActC), 96-hour microbial respiration assay (Resp [96 hr]), autoclaved citrate extractable protein index (ACE), wet aggregate stability (AGG), available water capacity (AWC), 24-hr microbial respiration (Resp [24 hr]), water-extractable organic C (WEOC), and water-extractable organic N (WEON) remained in one segment (0–15 cm), composited across a strip, and packed in coolers with ice packs for expedited shipping to soil analysis laboratories. Soil analyses at Ward Laboratories, Inc. (Kearney, NE) included loss-on-ignition OM, pH, P, and K. Soil samples in coolers were sent to Cornell University Soil Health Laboratory (Ithaca, NY) for the CASH suite analysis. Additional samples were sent to the USDA-Agriculture Research Service Grassland Soil Water Research Laboratory (Temple, TX) for the Haney Soil Health Tool (HSHT) suite of analyses (Haney et al., 2018). The USDA facility received samples from 2014 to 2018. Ward Labs completed the HSHT analyses during 2018 and 2019 from select locations only. A summary of soil test methods for each indicator is available in Table 3−1. Results were reported for the 0–15 cm depth layer to ensure comparability of nutrient and soil health test results. The two depths for the soil nutrient analyses were combined using weighted averages. The results from individual soil samples were averaged within a strip for each nutrient and soil health test.

Soil texture for each soil health sample was measured at the Cornell Soil Health Laboratory and classified using the USDA-NRCS system. Sand, silt, and clay observations reported by the Cornel Lab for each strip were averaged over time and were assigned to coarse (sand, loamy sand, and sandy loam), medium (loam, silt loam, silt), or fine (sandy clay loam, sandy clay, clay loam, silty clay loam, silty clay) texture groups (Figure 3−1). An official long-term climate normal value was challenging to obtain for every site; therefore, a long-term average for growing seasons (1 Apr to 31 Oct) between 1983 and 2013 was calculated for daily precipitation and daily maximum and minimum air temperature. Weather data were collected using the Daymet Single Pixel Extraction Tool (Oakridge National Laboratory, 2020) for the latitude and longitude corresponding to each site. Daymet pixels represented 1 km^2 of interpolated data. Cumulative growing degree days (GDD) were calculated using a ten-degree Celsius base temperature and 30°C as the crop maximum (North Dakota Agriculture Weather Network, 2019). Abundant and well-distributed rainfall (AWDR), a diversity measure calculated from daily rainfall, was used to describe the temporal availability of precipitation (Tremblay et al., 2012).

2.2 | Soil Health Indicator Laboratory Analyses

The suite of soil health indicators utilized in this analysis includes the most common soil health assessments in the U.S. and the most likely indicators used by growers and research practitioners in the Midwest, namely, the CASH, SMAF, and HSHT. Details of the current laboratory methods, protocols, and procedures used at the time of analysis for these soil samples were available from Ward Laboratories, Inc. (Ward Laboratories, 2020), Cornell University Soil Health Laboratory (Schindelbeck et al., 2016), and the USDA-ARS Grassland, Soil and Water Research Laboratory (Haney, 2020). A summary of the indicators, a description of their general purpose, and analysis methods are available in Table 1. Although some of the indicators measure similar soil properties (e.g., Resp [96 hr] and Resp [24 hr]), they are quantified by unique laboratory procedures in each soil health assessment (e.g., CASH and HSHT).

2.3 | Calculating Soil Health Indicator Assessment Scores

Soil health indicators for the CASH, SMAF, and HSHT were scored using methods described in Fine et al. (2017), Andrews et al. (2004), and Haney et al. (2018), respectively. The CASH scores were calculated using the SHP database as the scoring calibration dataset rather than the Cornell University Soil Health Lab database and are referred to as CASHSHP in this manuscript. The mean and standard deviation of the indicator observed values are included herein for comparison with other published values. In the CASH framework, OM, ActC, ACE, Resp (96 hr), AGG, and AWC were scored using the cumulative normal distribution function with the mean and standard deviation of three soil texture groups (Figure 1) (Fine et al., 2017). There were approximately 11%, 65%, and 24% of the observations in the coarse, medium, and fine texture groups, respectively. Therefore, before calculating the indicator scores, distribution normality was evaluated for each indicator within each soil texture group using Q-Q plots and histograms. Following evaluation, AGG lab values for coarse and medium soils were log-transformed. Additionally, ACE values were square root transformed for medium texture soils. pH scores were calculated according to threshold values available in Fine et al. (2017). According to the CASH manual (Moebius-Clune et al., 2016), nutrient element scores are based on local recommendations with threshold values. Therefore, CASH-scored macro and micronutrients were omitted from this analysis rather than formulating each nutrient's scoring thresholds by state or sub-state regions.

The SMAF scores were calculated using algorithms and logic statements parameterized by location-specific environment factors such as climate, soil suborder, field slope, and soil texture (Andrews et al. 2004). The OM loss on ignition was converted to soil organic carbon (SOC) using the conversion factor 0.58 described by Cambardella et al. (2001). The location-specific factors were identified using the USDA-NRCS Web Soil Survey area of interest tool (Soil Survey Staff, 2019) for each location. It is important to note that although the SMAF AGG score was originally parameterized from a different aggregate stability analysis method than the one used at the Cornell University Soil Health Lab, this study scored the Cornell AGG test results using the $CASH_{SHP}$ and the SMAF scoring approaches to evaluate the scores for all available indicators. The $CASH_{SHP}$ and SMAF AGG scores were then subjected to regression analysis to determine the agreement between the methods (see the Results and Discussion section). The HSHT scores were calculated according to Supplementary Equation 3−S1 available in the Supplementary Material. The $CASH_{SHP}$ and SMAF composite scores were an unweighted average of the individual indicator scores.

2.4 | Analytical Approach

The effect of soil texture on soil health indicator observed values was evaluated utilizing a repeated-measures analysis of variance (RMANOVA) linear mixed model in the GLIMMIX procedure of SAS (SAS Institute Inc., 2020). The soil texture group was considered a fixed effect, and the year was a random effect. Location nested within texture group was considered the subject of the repeated measures. The best fitting RMANOVA covariance structure for each indicator was selected by testing eight covariance structures and identifying the covariance structure with the lowest corrected Akaike's information criterion (Hurvich & Tsai, 1989). This RMANOVA model methodology was also used to test the influence of texture groups on individual soil health indicators and the assessment composite scores. The ls means statement with the Tukey-Kramer adjustment for unbalanced designs (Kramer, 1956) was used to complete post hoc analyses of group means separation.

Relationships between indicator scores and composite scores were evaluated to understand how individual indicators influence the composite scores. Furthermore, relationships among composite scores were assessed to understand how similarly they score soil health. The REG procedure of SAS was used to regress a composite assessment score on an individual indicator score in the $CASH_{SHP}$ and SMAF in an independent model for each soil texture group. Variation explained, represented by the adjusted R^2 value, was used as the evaluation criteria for determining an individual indicator's influence on the composite score. The HSHT was regressed on its indicator lab values because the HSHT does not score individual indicators. This analytic approach was used to quantify the variation explained in the composite score by the individual indicators to determine which indicators strongly influence a composite soil health score. Furthermore, simple linear regression with only one soil health indicator as the explanatory variable was used to avoid complications of multicollinearity among the indicators in a multiple-regression model. The CORR procedure of SAS was used to calculate Pearson's correlation coefficients among assessment composite scores for each soil texture group. Furthermore, the REG procedure was used to determine the amount of variation explained in the $CASH_{SHP}$ AGG scores by the SMAF AGG scores. Similarly, the relationships between the $CASH_{SHP}$ ACE and SMAF AWC scores were evaluated.

To understand whether corn and soybean yield was related to soil health scores, a regression approach was taken at three spatial and temporal scales, namely, 1) individual regression models were implemented for each site-year; thus, each site-year regression model was composed of four observations, 2) combined site-years at sites that had two or more years of data, resulting in site-level regression models having eight to 16 observations each, and 3) all site-years combined into a single regression model for each indicator. A custom SAS macro was used to identify significant P values by filtering SAS Output Delivery System tables from the REG procedure. Once significant models were identified, the mean adjusted R^2, and the frequency of a positive coefficient were calculated for each indicator using the FREQ procedure of SAS.

Following identification of site-years with significant yield-score regression models, it was hypothesized that a pattern of site-year environment factors might emerge. Thus, a binary variable representing responsiveness was used as the response variable in the LOGISTIC procedure of SAS with stepwise selection to fit a multiple logistic regression model to identify location factors and soil health indicators that explain the site-year associations between yield and a score. Significant yield-score regression models were labeled as responsive (1) or non-responsive (0), and site environment factors and soil health scores were used as independent variables in the logistic regression model. Mean site-year soil health indicator lab values were used in addition to silt and clay concentrations, tillage intensity (no-till, vertical-till, strip-till, and conventional-till) labeled as dummy variables, long-term average abundant and well-distributed rainfall (Tremblay et al., 2012), and growing degree days, and crop type (soybean or corn) as independent variables. Further details on validating the logistic model are available in the Supplemental Materials.

3 | RESULTS AND DISCUSSION

3.1 | Soil Health Indicator Observed Values and Assessment Scores

Soil health was assessed at 389 unique strips across 96 SHP locations using twelve soil health indicators common to the CASH, SMAF, and HSHT. In addition, there were 25 locations with one year of soil health data, 59 with two, 11 with three years, and three locations with four years of data over five years (Figure 3−1). Overall, the measured indicator values at these sites were similar (had overlapping ranges) to previous reports of indicator values for the Midwestern U.S. (Fine et al., 2017).

The repeated measures analysis of variance revealed that the means of eight of the twelve soil health indicators were different among soil texture groups (Table 3−2). The soil protein index (ACE), WEON, pH, and K were not influenced by soil texture across the SHP sites. In addition, indicator values generally increased as texture became finer in six of the eight indicators that varied among texture groups (Table 3−2). Additionally, the indicators' coefficients of variation (CV) were generally lower in the fine texture group (Supplementary Table 3−S1). However, AGG is a notable departure from these trends. Specifically, AGG lab values were lowest, and CV was highest in medium-textured soils. While the high CV may result from a large sample size, the lower observed value is perplexing because higher AGG values are often correlated with higher OM levels (Bronick & Lal, 2005). Furthermore, AGG has been shown to decrease with increasing soil tillage intensity (Weidhuner et al., 2021), yet 84% of the SHP locations in the medium texture group practiced either no-till or a form of reduced tillage (vertical or strip tillage). However, Fine et al. (2017) also observed lower AGG values in medium texture soils in the Midwestern U.S.

The $CASH_{SHP}$ approach scores indicators from a calibration dataset of soils only grouped by soil textural class rather than by region and texture as the CASH calculated at the Cornell University Soil Health Laboratory. Soil health assessment scores for individual indicators in the $CASH_{SHP}$ and SMAF reflected the trends detected in the observed values. When calculating scores in the $CASH_{SHP}$ framework, the mean observed value receives a score of 50 within texture groups (Moebius-Clune et al., 2017). Consequently, the mean $CASH_{SHP}$ indicator scores were not different among texture groups. However, the $CASH_{SHP}$ indicator curves (Figure 2) reflect the distributions of observed values (Table 3−2; Supplementary Table 3−S1).

The SMAF parameterizes scoring curves based on soil taxonomy, texture, climate, topographical slope, and threshold values. Therefore, the SMAF scores create different curves according to the algorithms' factor-level classes (Figure 3−3). For example, soils within the same texture group but from different climatic zones will be scored differently; thus, the scores along the same curve represent similar soils from the same climatic zone. The mean SMAF indicator scores for SOC, K, AGG, and AWC varied among soil texture groups (Table 3−3). In contrast to the CASH and SMAF, the HSHT only uses its indicator threshold values to calculate an overall soil health score (Supplementary Equation 3−1); the individual soil health indicators are not scored in the HSHT.

3.2. | Soil Health Assessment Composite Scores

3.2.1 | Summary of assessment composite scores

Assessment composite scores provide an overall evaluation of a soil's health and guide soil health comparisons (Moebius-Clune et al., 2017). The $CASH_{SHP}$ and SMAF composite scores are un-weighted means of their indicator scores, while the HSHT is a weighted computation of Resp (24 hr), WEOC, and WEON. Except for the $CASH_{SHP}$, the assessment composite scores increased as texture became finer (Supplementary Table 3−S2). Specifically, the mean $CASH_{SHP}$ composite score was 52.7 to 55.6 across texture groups. Soil Management Assessment Framework composite score means were 7.6, 8.2, and 8.6 for coarse, medium, and fine textures, respectively, and were different at the 0.001 probability level. The HSHT mean score for coarse and medium soils was 13.55. The average score for fine textured soils was 16.3 and was different from the coarse and medium texture mean score at the 0.05 probability level; however, the means of the coarse and medium soils were not different. Although, it should be noted that those two texture groups had smaller sample sizes than the medium texture group. The composite score standard errors of the means are available in Supplementary Table 3−S2.

Correlations among the soil health assessment composite scores may demonstrate the level of agreement among the assessments when scoring a soil. The SMAF and $CASH_{SHP}$ appeared to have the most robust agreement among the three assessments, albeit only in medium and fine-textured soils (Table 3−4). The HSHT had consistently weaker correlations with the CASH and SMAF. The four indicators shared between the CASH and SMAF might plausibly explain the strong correlations between the two composite scores. However, the HSHT and CASH share only one similar indicator, a measure of microbial respiration. That similarity might explain the slightly stronger correlations between the $CASH_{SHP}$ and HSHT than with the SMAF composite score. Other possible sources of agreement among the SMAF and CASH versus the HSHT composite scores might be how composite scores are calculated. Furthermore, across all texture groups, a simple linear regression of $CASH_{SHP}$ AGG scores regressed on the SMAF AGG scores demonstrated strong agreement ($R^2 = 0.83$). However, the AWC scores for $CASH_{SHP}$ and SMAF were moderately related ($R^2 = 0.45$).

Each of these assessment methods has strengths and weaknesses. For example, grouping soils into categories is fundamentally the discretization of continuous variables. As shown in the right panel of Figure 3−1, the mean sand, silt, and clay concentrations in the coarse group are primarily separated from the texture concentrations of the medium and fine texture groups. However, there is no clear distinction between fine and medium-textured soils. Thus, thresholds defining soil texture groupings might be most appropriate in certain conditions where the assessments were developed (i.e., Midwest U.S.) and may require thorough testing before those thresholds are generally applied across soil types, geographies, and cropping systems. However, discretization makes it possible to calculate a cumulative normal curve or parameterize algorithms with fewer observations to generate a suitable scoring curve. With these drawbacks in mind, a recent study demonstrated that applying soil survey data to unsupervised machine learning can classify soils to reflect natural differences in soil properties and characteristics in the Western U.S. (Devine et al., 2020). However, this technique was not applied to calibrating soil health assessment scores. Consequently, additional work is needed to evaluate the effectiveness of these soil health score calibration methods across many geographies.

3.2.2 | Influence of indicators on composite scores

Individual regression analyses were completed for the $CASH_{SHP}$, SMAF, and HSHT composite scores by independently regressing each composite score on each of their indicator scores (lab values for the HSHT) by soil texture group. Active carbon, OM, and ACE independently accounted for approximately 50% of the variation within their respective regression models for $CASH_{SHP}$ when averaged across texture groups (Figure 3−3). In the SMAF, AGG and SOC more consistently accounted for variation in the SMAF composite score across soil texture groups than other SMAF indicators (Figure 3−3). In the HSHT, Resp (24 hr) consistently accounted for large proportions of the HSHT score variation across soil texture groups. One explanation for this might be how the HSHT weights Resp (24 hr) differently than WEOC and WEON. However, WEOC and WEON accounted for >50% of the HSHT score variation in their respective regression models for coarse and medium texture soils. Across the assessments, soil texture had a distinct effect on how a composite score responded to each indicator (Figure 3−3). Additionally, biological indicators ActC, OM/SOC, ACE, microbial respiration measures, WEOC and WEON, and the physical indicator AGG greatly affected their composite scores.

Overall, the $CASH_{SHP}$ and SMAF composite scores were particularly influenced by the behavior of indicator scores according to soil texture. Specifically, an indicator's score may have a large effect on the composite score in one texture group but not others. Additionally, within texture groups across indicators, the composite scores' variation was not accounted for equally by all the indicators (Figure 3−3). This is problematic for soil health score interpretation and may suggest that other composite score calculation methods that account for soil properties differentially may be more appropriate. For example, Congreves et al. (2015) reported that when principal components analysis was used to weight the indicators by their principal component loadings, the $CASH_{SHP}$ score could better identify differences in soil health management practices. However, before the SMAF was formally introduced, an extensive review of composite score calculation methods was undertaken to determine if an unweighted or weighted average was most appropriate. In that study, Andrews et al. (2002) decided that an unweighted calculation method was sufficient to characterize soil health in a composite assessment score. Results from the present study demonstrated that additional work on soil health composite scores might be necessary.

3.3 | Yield and Indicator Scores

3.3.1 | Corn and soybean yield correlations with soil health assessment scores

Three spatial and temporal scales of analysis were used to evaluate how soil health indicator and assessment composite scores relate to crop yield in the control strips at Soil Health Partnership sites. For each soil health indicator and assessment composite score, the relationship between scores and yield was first analyzed on an individual site-year basis. Then, an additional analysis combined multiple years of data at each site when there were more than two years of data at a location. Finally, all sites and years were analyzed together.

There were 155 site years and 56 sites available for analysis comprised of 84 site-years of corn and 71 site-years of soybean yields. Less than eight percent of the 155-independent site-years were significant when considering each indicator and composite scores independently. However, altogether, there were 72 unique site-years (46%) with a responsive regression model. Thirty-three site-years (21%) had responsive models with at least two indicators or composite scores. In the individual site analysis, the yield was related to an indicator or composite score at least once at all 56 sites, and 44 unique sites

(78%) had more than two responsive yield-score relationships. When all data were combined, 11 of 16 indicators (69%) had a significant response. The amount of variation explained in crop yield by soil health scores was higher for individual site-years than for individual sites, where the mean adjusted R^2 was 0.92 versus 0.65, respectively (Figure 3). However, at the broadest scale of analysis, the mean adjusted R^2 value was 0.03 (Supplemental Table 3−S3). For example, observations within the same site-year were closer to each other than observations from one year to the next, most likely due to annual oscillations in weather and management that dynamically influence crop production and these soil health indicators. The $CASH_{SHP}$, SMAF, and HSHT each explained similar amounts of variation in corn or soybean yield (Figure 3−3). The mean frequency of observing a positive regression slope among the site-year scale models was 65% for SMAF indicators and 51% for $CASH_{SHP}$ indicators. At the individual site level, the mean frequency for observing a positive slope among $CASH_{SHP}$ indicator models was 53% and 45% for SMAF indicators. Notably, among composite scores, the lowest frequency of positive regression slopes (36%) occurred in the $CASH_{SHP}$ composite score at the site-year scale. In contrast, other composite scores, and at different scales, ranged between 47 and 82% (Figure 3−3). With all data combined at the broadest scale of analysis, 87% of the responsive soil health scores had a positive slope (Supplemental Table 3−S3). Overall, there was a greater number of responsive scores from $CASH_{SHP}$ indicators among the three scales of analysis than SMAF indicators (Figure 3−3).

3.3.2 | Explaining associations between yield and scores with multiple logistic regression

Multiple logistic regression with stepwise selection was utilized to fit a probability prediction model of site-year correlations between yield and soil health assessment indicator and composite scores. A binary response variable was used to represent responsive site-year models, while location condition factors (Supplemental Table 3−S1) and soil health scores were utilized as predictor variables in the multiple logistic regression model. It was hypothesized that interactions among location environment factors, soil functions, and crop yield would be detectible in a pattern across the study area. The model revealed that none of the independent variables could predict a responsive relationship between yield and an assessment score for any site-year. Unfortunately, without any apparent patterns to explain why responsive site-years were observed, this result showed that factors leading to a relationship between yield and a soil health assessment score were not consistent among observations from 155 site-years. Future work might employ methods that more fully account for system complexity, such as structural equation modeling (Wade et al., 2020), that may offer greater insight than the forgoing logistic regression analysis.

These results demonstrated that interpreting the relationship between a soil health score and crop yield depends on the indicator and location, among many other factors not yet accounted for in these simple models. Furthermore, these results demonstrated a concept of soil health assessment: As the temporal and spatial scale expands, the concomitant increase in variability provides evidence that soil health assessment is predominantly a localized endeavor. For example, when multiple years of data were analyzed at each site, the proportion of significant models increased (Figure 3−3B). However, when all sites and site-years were combined at the broadest scale of analysis, the coefficient of determination was the lowest. Essentially, the additional data may increase the detection of a statistical correlation between yield and a soil health score. Still, due to location-specific and annual environment factors, additional data did not guarantee that the correlation would be a strong one. Furthermore, including other factors to account for year-to-year variation when analyzing the relationship between crop production and soil health scores may be needed in future analyses. For example, in a previous study, Crookston et al. (2021) used soil health indicator lab values and location-specific factors, such as the long-term climate average and field tillage, in a multiple regression model estimating corn and soybean yield at SHP sites. Crookston et al. (2021) reported that soil health indicators had a weak influence on yield relative to the other variables in the model. Together, these results indicate that the direct impact of soil health indicators on crop production is difficult to disentangle from environmental factors (climate and management). Specifically, the potential utility of soil health testing is not diminished because of the lack of strong relationships with yield. These results may encourage practitioners to measure additional ecosystem service and environmental outcomes related to soil health functions and processes.

4 | CONCLUSIONS

These analyses of soil health assessment frameworks identified many challenges facing interpreting soil health scores for the $CASH_{SHP}$, SMAF, and HSHT. The assessment composite scores did not always agree, and the correlation strength among the scores was modified by soil texture. Further studies of soil health scores may also elaborate on the connection between soil health scores and soil health outcomes beyond crop yield, such as water quality or biodiversity. Caution is also warranted so that soil health monitoring is not dismissed as irrelevant because of the low frequency of correlations between soil health scores and crop yield. These results most strongly demonstrate that soil health monitoring is a process that requires commitment and consistency over many seasons to observe the relationships between soil health measurements and soil health outcomes. It must also be noted that these analyses utilized the non-treated strips of the Soil Health Partnership on-farm trials where only two to five years of data were available. Thus, long-term studies that periodically measure soil health are more relevant than ever in supporting soil health-based crop and soil management.

5 | TABLES AND FIGURES

TABLE 3‒1 Soil health indicator abbreviations, units, description, soil health assessment (SHA), and laboratory analysis methods used in this study.

Soil health indicator	Description	SHA[a]	Analysis method	Citation
Organic matter (OM); soil organic carbon (SOC)	Carbon-based materials originating from living organisms	CASH, SMAF	Calculated as weight lost from a soil sample on ignition. SOC was calculated by multiplying percent OM by 0.58.	Schindelbeck et al., 2016; Cambardella et al., 2001
Permanganate oxidizable carbon (active carbon) (ActC)	A measure of readily available organic carbon energy source for soil microbes.	CASH, SMAF	Photospectrometry analysis of oxidized potassium permanganate extractant.	Schindelbeck et al., 2016
Autoclaved citrate extractable soil protein index (ACE)	A measure of organically bound nitrogen. Microbial activity makes this organic matter fraction available for plant use.	CASH	High pressure and temperature extraction of citrate solution.	Schindelbeck et al., 2016
Soil microbial respiration 96-hour incubation (Resp (96 hr))	A measure of soil microbial metabolic activity.	CASH	Quantification of CO_2 gas trapped in solution evolved from re-wetted soil incubated 96 hours.	Schindelbeck et al., 2016
Soil microbial respiration 24-hour incubation (Resp (24 hr))	A short duration measure of soil microbial metabolic activity.	HSHT	Paper chromatography quantification of CO_2 gas evolved from re-wetted soil incubated 24 hours.	(Haney, 2020; Ward Laboratories, 2020)
Water-extractable organic carbon (WEOC)	A measure of readily available organic carbon energy source for soil microbes.	HSHT	Quantification of organic C extracted with water from a soil sample.	(Haney, 2020; Ward Laboratories, 2020)
Water-extractable organic nitrogen (WEON)	A measure of organically bound nitrogen. Considered as a "nutritional" source for microbes.	HSHT	Quantification of organic N extracted with water from a soil sample.	(Haney, 2020; Ward Laboratories, 2020)
pH	Affects the availability of nutrients and biological properties in the soil.	CASH, SMAF	Voltage meter calibrated to determine Hydrogen ion activity in soil solution.	Watson and Brown, 1998
Soil chemical nutrients: P, K	Soil nutrients are needed for healthy plant growth.	SMAF	Mehlich-III extractant method and quantified using inductively coupled atomic plasma spectroscopy.	Soil and Plant Analysis Council, 1999; Warncke & Brown, 1998
Available water capacity (AWC)	Soil water available for plant uptake.	CASH, SMAF	Amount of water extracted from a pulverized and sieved soil sample using a pressure chamber.	Schindelbeck et al., 2016
Soil wet aggregate stability (WAS)	The proportion of soil aggregates resistant to degradation following rain.	CASH, SMAF	Calculated from soil remaining on a 0.25 mm sieve following simulated rainfall.	Schindelbeck et al., 2016
Sand, silt and clay	Soil proportions of particle size 0.002—0.05 mm (silt) and less than 0.002 mm (clay).		Rapid 4-hour quantification of sand, silt, and clay from soil/water solution.	Schindelbeck et al., 2016

[a] Comprehensive Assessment of Soil Health (CASH), Soil Management Assessment Framework (SMAF), and the Haney Soil Health Tool (HSHT).

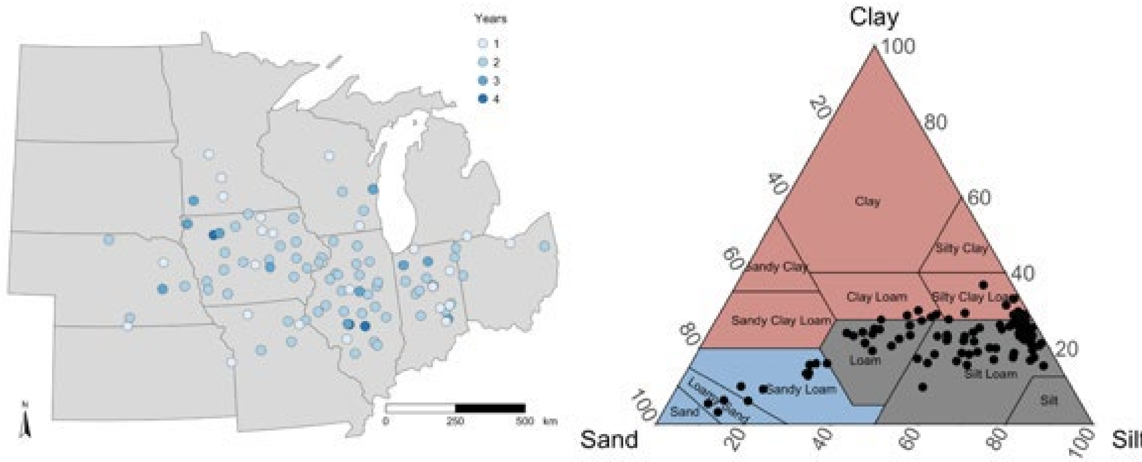

FIGURE 3−1 The number of data years at 96 Soil Health Partnership locations (left) and their mean soil texture (right). Soil texture groups were defined as coarse: sand, loamy sand, sandy loam; medium: sandy clay loam, loam, silt loam, silt; fine: sandy clay loam, clay loam, silty clay loam, silty clay, clay.

TABLE 3−2 Soil health indicator mean values by texture group, standard error of the mean (SEM), significant probability level ($P > F$). Significant mean separation (after Tukey adjustment) is indicated by different lowercase letters within each row. Texture groups with the same lowercase letter are not significantly different from each other. See Table 3−1 for variable definitions and Figure 3−1 for texture group descriptions

Soil health indicator	**Soil texture group**			**SEM**	***P > F*** [a]
	Coarse	**Medium**	**Fine**		
OM (g kg^{-1})	26.79*c*	32.30*b*	45.21*a*	2.82	***
ActC (mg kg^{-1})	406.07*c*	505.34*b*	616.18*a*	30.48	***
ACE (mg g^{-1})	5.28	4.89	5.28	0.24	ns
Resp (96 hr) (mg CO_2 C g^{-1})	0.41*c*	0.46*b*	0.50*a*	0.03	*
Resp (24 hr) (mg CO_2 C kg^{-1})	44.68*b*	52.56*b*	71.28*a*	10.18	*
WEOC (mg kg^{-1})	222.94*c*	232.8*b*	263.92*a*	17.11	*
WEON (mg kg^{-1})	21.01	20.82	22.87	1.52	ns
pH	6.37	6.48	6.48	0.1	ns
P (mg kg^{-1})	90.19*a*	41.64*b*	40.90*b*	10.24	***
K (mg kg^{-1})	165.82	184.63	197.76	22.47	ns
AGG (%)	27.32*a*	18.64*b*	29.47*a*	2.4	***
AWC (g g^{-1})	0.16*b*	0.29*a*	0.28*a*	0.01	***

[a] *Significant at the 0.05 probability level; **Significant at the 0.01 probability level; ***Significant at the 0.001 probability level; ns, not significant.

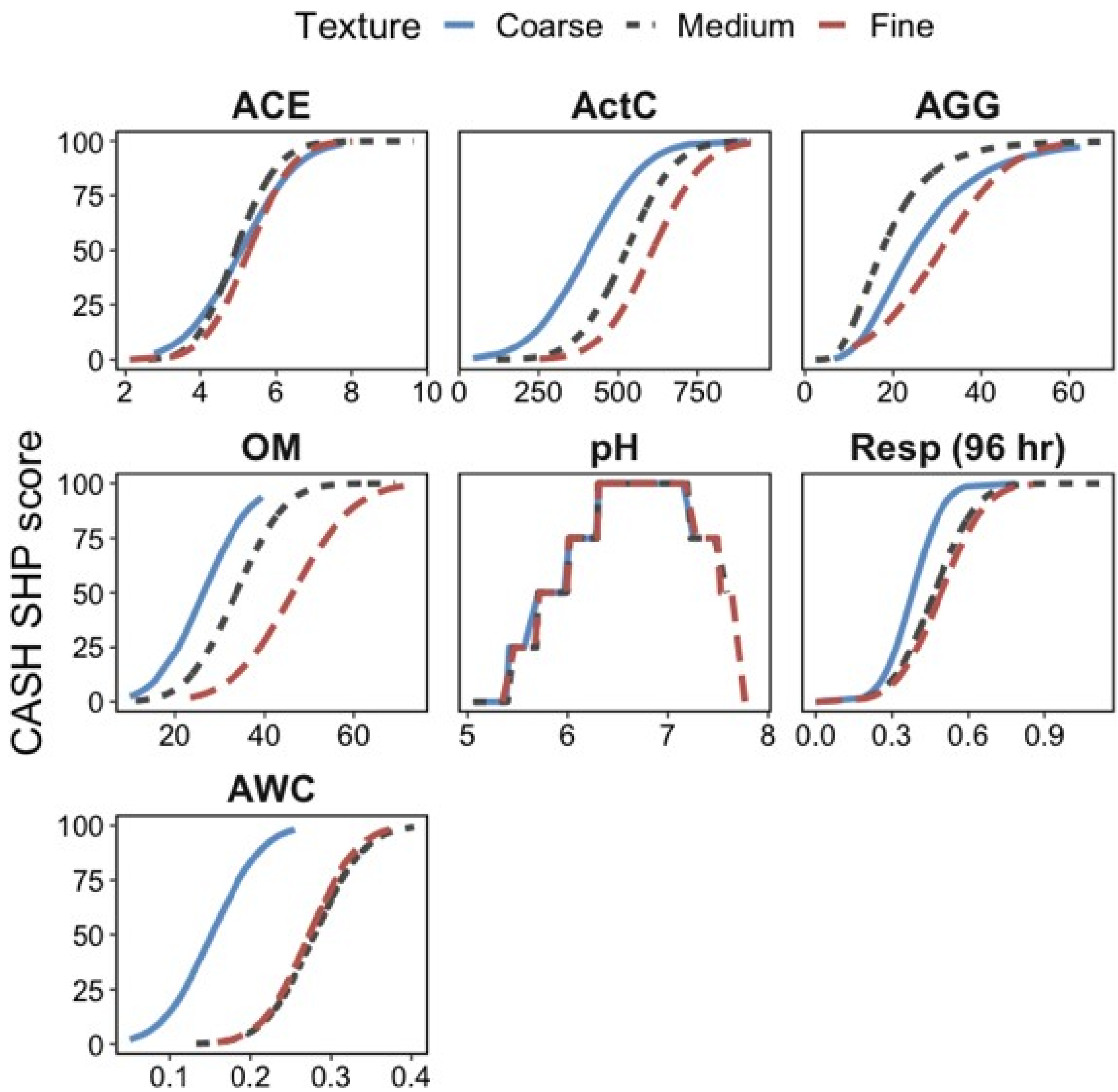

FIGURE 3‒2 Comprehensive Assessment of Soil Health indicator scores for Soil Health Partnership sites (CASH SHP). See Figure 3‒1 for texture group descriptions. ACE, autoclaved citrate extractable protein index (mg g^{-1}); ActC, active carbon (mg g^{-1}); AGG, wet aggregate stability (%); AWC, available water capacity (g g^{-1}); OM, organic matter loss on ignition (g kg^{-1}); Resp (96 hr), microbial respiration 96-hour incubation (mg CO_2 C g^{-1})

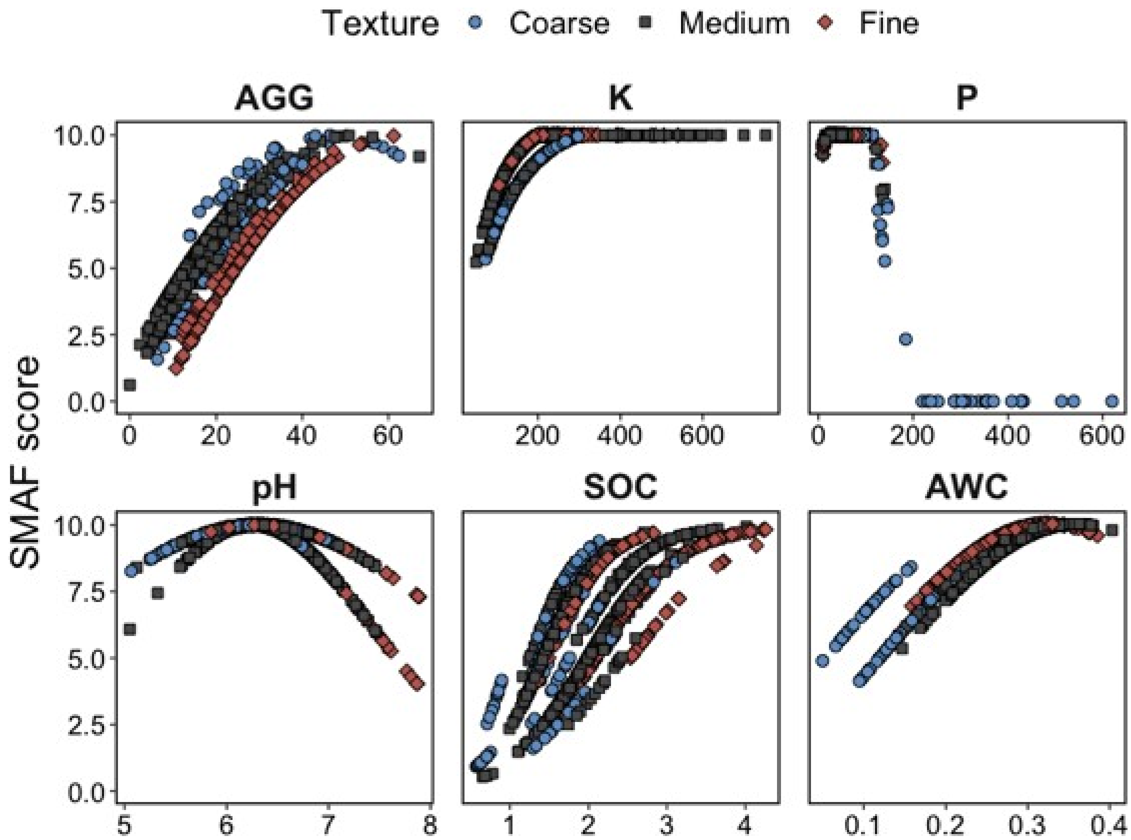

FIGURE 3‒3 Soil Management Assessment Framework (SMAF) scores for Soil Health Partnership. Separate curves within the same texture group affect soil factor classes used in the SMAF score calculations. See Figure 3‒1 for texture group descriptions. AGG, wet aggregate stability (%); AWC, available water capacity (g g^{-1}); K, potassium (mg kg^{-1}); P, phosphorus (mg kg^{-1}); SOC, soil organic carbon (%).

TABLE 3−3 Mean Soil Management Assessment Framework (SMAF) soil health indicator scores by soil texture group, standard error of the mean (SEM), and statistical significance (*P* > *F*). Significant mean separation (after Tukey adjustment) is indicated by different lowercase letters within each row. Texture groups with the same lowercase letter are not significantly different from each other. See Table 3−1 for variable definitions and Figure 3−1 for texture group descriptions

	SMAF indicator score (0–10)				
Soil health indicator	Soil texture group			SEM	*P* > *F*[a]
	Coarse	Medium	Fine		
SOC	5.11*b*	5.72*b*	7.47*a*	0.60	***
pH	9.57	9.64	9.71	0.08	ns
P	9.53	9.27	9.69	0.51	ns
K	8.19*c*	9.15*b*	9.63*a*	0.24	***
AGG	7.35*a*	5.52*b*	6.10*ab*	0.50	***
AWC	7.01*b*	9.42*a*	9.67*a*	0.12	***

[a] ***Significant at the 0.001 probability level; ns, not significant.

TABLE 3−4 Pearson correlation coefficients ($P < 0.05$) by soil texture group for the Comprehensive Assessment of Soil Health for the Soil Health Partnership ($CASH_{SHP}$), Soil Management Assessment Framework (SMAF), Haney Soil Health Tool (HSHT). See Figure 3−1 for texture group descriptions.

Texture group		**CASH**	**SMAF**
Coarse			
	SMAF	ns[a]	
	HSHT	0.45	ns
Medium			
	SMAF	0.64	
	HSHT	ns	0.22
Fine			
	SMAF	0.73	
	HSHT	0.34	0.21

[a] ns, not significant.

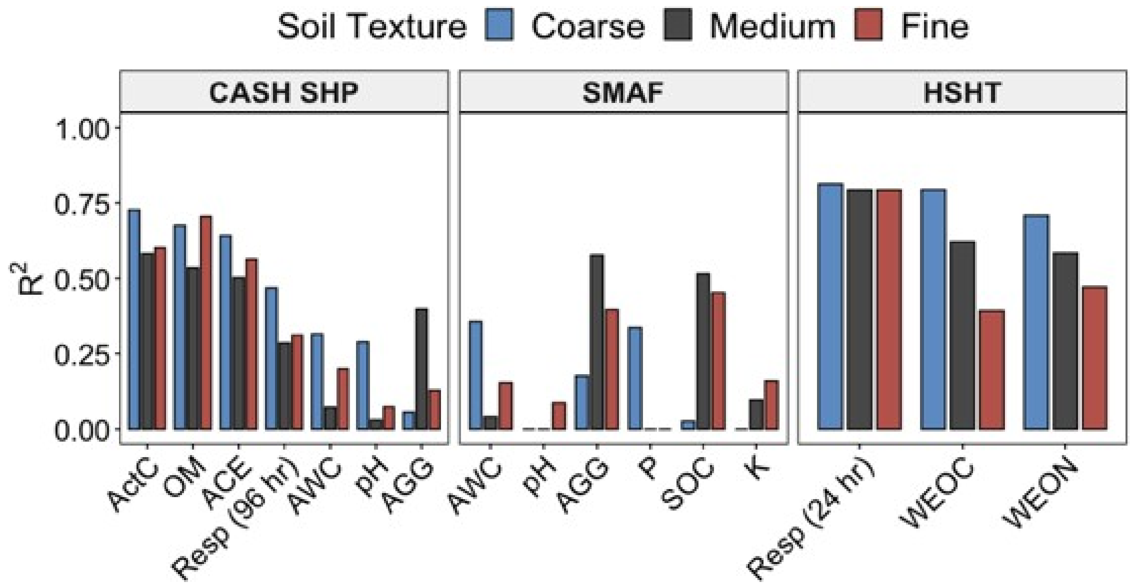

FIGURE 3‒4 Composite score variation explained (R^2) by the scores of individual soil health indicators of the Comprehensive Assessment of Soil Health for the Soil Health Partnership ($CASH_{SHP}$), Soil Management Assessment Framework (SMAF), and observed values of the Haney Soil Health Tool (HSHT) grouped by soil texture. See Figure 3‒1 for texture group descriptions and Table 3‒1 for variable definitions

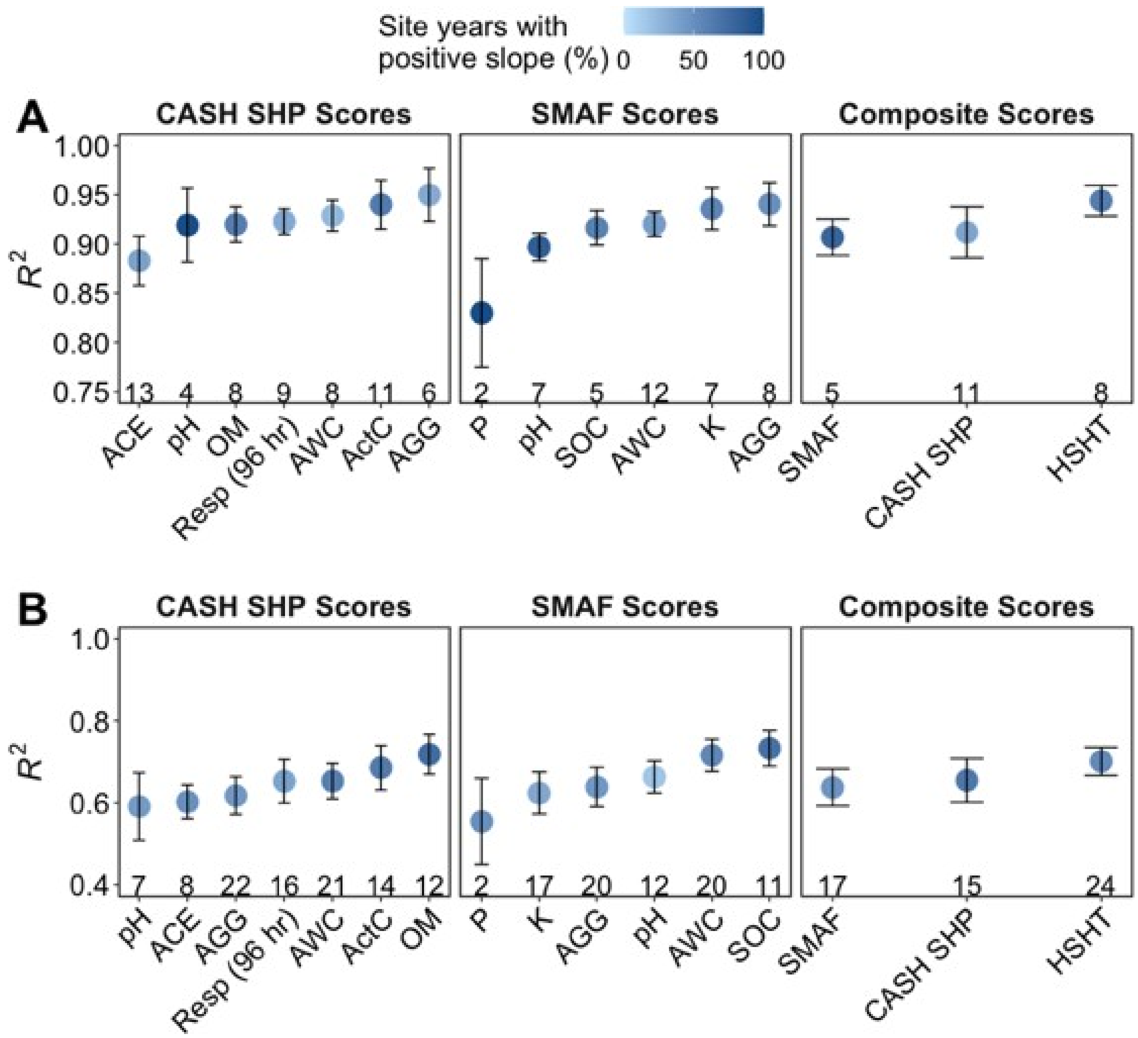

FIGURE 3‒5 Mean coefficient of determination (R^2) for corn and soybean yield regressed on individual indicator (Comprehensive Assessment of Soil Health for Soil Health Partnership sites [$CASH_{SHP}$], Soil Management Assessment Framework [SMAF]) and composite scores (SMAF, CASH, and Haney Soil Health Tool [HSHT]). Results are presented for individual site-year (A) and individual site (B) analyses. Error bars indicate the standard error of the mean; *n* is the number of site years (A) or sites (B) with a significant regression model.

6 | SUPPLEMENTARY MATERIAL

Equation 3‒S1

When WEOC:WEON is < 5, then

$$HHMMHHHH = \frac{RRRRRRRR\ (24\ hr)}{10} + \frac{WWMMWWWW}{50} + \frac{WWMMWWWW}{100}$$

If Resp (24 hr) < 100 mg CO_2 C kg^{-1} then,

$$HHMMHHHH = \frac{RRRRRRRR\ (24\ hr)}{10} + \frac{WWMMWWWW}{50} + \frac{WWMMWWWW}{10}$$

if Resp (24 hr) is 100–200 mg CO_2 C kg^{-1} then,

$$HHMMHHHH = \frac{RRRRRRRR\ (24\ hr)}{12} + \frac{WWMMWWWW}{50} + \frac{WWMMWWWW}{10}$$

if Resp (24 hr) is 200–300 mg CO_2 C kg^{-1} then,

$$HHMMHHHH = \frac{RRRRRRRR\ (24\ hr)}{14} + \frac{WWMMWWWW}{50} + \frac{WWMMWWWW}{10}$$

if Resp (24 hr) is 300–400 mg CO_2 C kg^{-1} then,

$$HHMMHHHH = \frac{RRRRRRRR\ (24\ hr)}{16} + \frac{WWMMWWWW}{50} + \frac{WWMMWWWW}{10}$$

if Resp (24 hr) is 400–500 mg CO_2 C kg^{-1} then,

$$HHMMHHHH = \frac{RRRRRRRR\ (24\ hr)}{18} + \frac{WWMMWWWW}{50} + \frac{WWMMWWWW}{10}$$

if Resp (24 hr) is >500 mg CO_2 C kg^{-1} then,

$$HHMMHHHH = \frac{RRRRRRRR\ (24\ hr)}{20} + \frac{WWMMWWWW}{50} + \frac{WWMMWWWW}{10}$$

where Resp (24 hr) is soil microbial respiration in 24-hour incubation, WEOC is water-extractable carbon, WEON is water-extractable nitrogen.

Table 3−S1 Sample size (n) and median (Med), minimum (Min), Maximum (Max), and coefficient of variation (CV) for 12 soil health indicators by soil texture group.

Texture[a]	Indicator	n	Med	Min	Max	CV
Coarse	AGG (%)	75	27.09	6.43	62.50	42.34
	AWC (g g^{-1})	75	0.14	0.07	0.26	31.72
	OM (g kg^{-1})	75	26.67	10.17	53.67	31.05
	ActC (mg kg^{-1})	75	442.51	122.79	905.00	32.19
	ACE (mg g^{-1})	75	5.38	2.76	10.14	25.54
	Resp (96 hr) (mg CO_2 C g^{-1})	75	0.39	0.00	0.79	28.37
	pH	75	6.47	5.27	7.26	7.04
	P (mg kg^{-1})	75	60.11	23.29	619.87	107.10
	K (mg kg^{-1})	75	168.56	71.93	344.92	34.43
	Resp (24 hr) (mg CO_2 C kg^{-1})	58	55.70	16.20	160.00	51.17
	WEOC (mg kg^{-1})	58	205.90	61.00	353.90	39.89
	WEON (mg kg^{-1})	58	19.60	4.50	34.00	40.46
Medium						
	AGG	491	18.08	2.37	67.18	45.31
	AWC	491	0.28	0.13	0.40	17.57
	OM	438	33.75	11.25	69.20	25.71
	ActC	490	514.59	117.85	918.40	23.03
	ACE	491	4.92	2.60	9.65	16.36
	Resp (96 hr)	491	0.44	0.19	1.12	27.31
	pH	438	6.51	5.12	7.47	5.86
	P	438	36.96	14.16	139.75	46.83
	K	438	170.33	66.83	753.33	48.62
	Resp (24 hr)	443	58.90	10.70	465.00	80.52
	WEOC	443	221.00	50.00	584.00	38.81
	WEON	443	19.80	4.70	48.90	38.63
Fine						
	AGG	186	28.23	10.79	61.23	35.49
	AWC	186	0.28	0.16	0.39	16.59
	OM	174	42.58	23.44	72.89	23.72
	ActC	186	612.65	253.28	937.19	22.16
	ACE	185	5.25	2.12	8.02	18.69
	Resp (96 hr)	186	0.47	0.00	1.74	31.19
	pH	174	6.46	5.47	7.88	6.90
	P	174	36.67	11.67	132.67	56.41
	K	174	185.50	81.40	539.72	37.85
	Resp (24 hr)	170	78.40	8.70	297.70	59.21
	WEOC	170	241.25	89.00	510.00	33.03
	WEON	170	20.80	7.30	36.00	30.91

[a] coarse (sand, loamy sand, and sandy loam); medium (loam, silt loam, silt); fine (sandy clay loam, sandy clay, clay loam, silty clay loam, silty clay).

Table 3−S2 Mean composite scores for the Comprehensive Assessment of Soil Health for Soil Health Partnership sites ($CASH_{SHP}$), Soil Management Assessment Framework (SMAF), Haney Soil Health Tool (HSHT). Standard error of the mean (SEM), and statistical significance ($P > F$), and covariance structure (COV). Mean separation is indicated by different lowercase letters.

Assessment	Soil texture group			SEM	$P > F$[a]
	Coarse	Medium	Fine		
$CASH_{SHP}$ (0–100)[b]	52.74	54.64	55.59	3.62	ns
SMAF (0–100)	7.57*c*	8.2*b*	8.64*a*	0.16	***
HSHT (0–50)	13.05*b*	13.99*b*	16.29*a*	1.39	*

[a] *Significant at the 0.05 probability level; ***Significant at the 0.001 probability level; ns, not significant.
[b] Number in parentheses indicates the minimum and maximum possible score.

7 | SUPPLEMENTAL METHODS. Validating multiple logistic regression

Before fitting the logistic model, the dataset of 155 site-years was randomly split (70/30) into training and validation datasets. The CTABLE option was specified in the model statement of the LOGISTIC procedure of SAS to call an Output Delivery System table of probability levels corresponding to correct classification frequencies, where the critical probability level was identified that maximized the percentage of correctly classified responsive site-years (sensitivity) and non-responsive site-years (specificity) (Allison, 2012). The store statement was invoked to output a table of the selected parameters and their coefficients to validate the model. The model out-table was then restored in the PLM procedure of SAS to score the validation dataset (SAS Institute Inc., 2020). The predicted probabilities scored in the PLM procedure were used in a SAS DATA if/then statement to classify site-years as responsive or non-responsive according to the critical probability level identified by the training model such that all location-years with a predicted probability above the critical level were classified as responsive. The model fit was evaluated using the concordant/discordant pairs and Somers' D values in the training set and the FREQ procedure of SAS to identify the frequency of correctly classified pairs in the validation set.

8 | REFERENCES

Andrews, S. S., Karlen, D. L., & Cambardella, C. A. (2004). The Soil Management Assessment Framework. *Soil Sci. Soc. Am. J.*, *68*(6), 1945–1962. https://doi.org/10.2136/sssaj2004.1945

Andrews, S. S., Karlen, D. L., & Mitchell, J. P. (2002). A comparison of soil quality indexing methods for vegetable production systems in Northern California. *Agriculture, Ecosystems and Environment*, *90*(1), 25–45. https://doi.org/10.1016/S0167-8809(01)00174-8

Bronick, C. J., & Lal, R. (2005). Soil structure and management: A review. *Geoderma*, *124*(1–2), 3–22. https://doi.org/10.1016/j.geoderma.2004.03.005

Cambardella, C. A., Gajda, A. M., Doran, J. W., Wienhold, B. J., & Kettler, T. A. (2001). Estimation of Particulate and Total Organic Matter by Weight Loss-on-Ignition. In L. R., K. J.M., F. R.F., & S. B.A. (Eds.), *Assessment Methods for Soil Carbon* (pp. 349–359). CRC Press.

Cherubin, M. R., Karlen, D. L., Cerri, C. E. P., Franco, A. L. C., Tormena, C. A., Davies, C. A., & Cerri, C. C. (2016). Soil quality indexing strategies for evaluating sugarcane expansion in Brazil. *PLoS ONE*, *11*(3), 1–26. https://doi.org/10.1371/journal.pone.0150860

Congreves, K. A., Hayes, A., Verhallen, E. A., & Van Eerd, L. L. (2015). Long-term impact of tillage and crop rotation on soil health at four temperate agroecosystems. *Soil and Tillage Research*, *152*, 17–28. https://doi.org/10.1016/j.still.2015.03.012

Crookston, B. S., Yost, M. A., Bowman, M., Veum, K., Cardon, G., & Norton, J. (2021). Soil health spatial-temporal variation influence soil security on Midwestern, U.S. farms. *Soil Security*, *3*. https://doi.org/10.1016/j.soisec.2021.100005

Devine, S. M., Steenwerth, K. L., & O'Green, A. T. (2020). A Regional Soil Classification Framework to Improve Soil Health Diagnosis and Management. *Soil Sci. Soc. Am. J.*, *85*, 361–378. https://doi.org/10.1002/saj2.20200.

Doran, J. W., & Zeiss, M. R. (2000). Soil health and sustainability: managing the biotic component of soil quality. *Applied Soil Ecology*, *15*(1), 3–11. https://doi.org/10.1016/S0929-1393(00)00067-6

Fine, A. K., van Es, H. M., & Schindelbeck, R. R. (2017). Statistics, Scoring Functions, and Regional Analysis of a Comprehensive Soil Health Database. *Soil Sci. Soc. Am. J.*, *81*, 589–601. https://doi.org/10.2136/sssaj2016.09.0286

Hammac, W. A., Stott, D. E., Karlen, D. L., & Cambardella, C. A. (2016). Crop, Tillage, and Landscape Effects on Near-Surface Soil Quality Indices in Indiana. *Soil Science Society of America Journal*, *80*(6), 1638. https://doi.org/10.2136/sssaj2016.09.0282

Haney, R. L. (2020). *Soil Health Tool (SHT) version 4.7*. https://www.nrcs.usda.gov/wps/PA_NRCSConsumption/download/?cid=nrcseprd338485&ext=pdf

Haney, R. L., Haney, E. B., Smith, D. R., Harmel, R. D., & White, M. J. (2018). The soil health tool—Theory and initial broad-scale application. *Applied Soil Ecology*, *125*, 162–168. https://doi.org/10.1016/j.apsoil.2017.07.035

Harmel, R. D., & Haney, R. L. (2013). Initial Field Evaluation of the Agro-Economic Effects of Determining Nitrogen Fertilizer Rates with a Recently-Developed Soil Test Methodology. *Open Journal of Soil Science*, *03*(02), 91–99.

https://doi.org/10.4236/ojss.2013.32010
Hurvich, C. M., & Tsai, C.-L. (1989). Regression and Time Series Model Selection in Small Samples. *Biometrika*, *76*(2), 297–307. https://doi.org/https://doi.org/10.1093/biomet/76.2.297
Karlen, D. L., Mausbach, M. J., Doran, J. W., Cline, R. G., Harris, R. F., & Schuman, G. E. (1997). Soil Quality: A Concept, Definition, and Framework for Evaluation (A Guest Editorial). *Soil Science Society of America Journal*, *61*(1), 4. https://doi.org/10.2136/sssaj1997.03615995006100010001x
Kramer, C. Y. (1956). Extension of Multiple Range Tests to Group Means with Unequal Numbers of Replications. *Biometrics*, *12*, 307–310. https://doi.org/https://doi.org/10.2307/3001469
Kyveryga, P. M., Mueller, T. A., & Mueller, D. S. (2018). On-Farm Replicated Strip Trials. In D. K. Shannon, D. E. Clay, & N. R. Kitchen (Eds.), *Precision Agriculture Basics* (pp. 189–207). American Society of Agronomy, Crop Science Society of America, and Soil Science Socitey of America. https://doi.org/10.2134/precisionagbasics.2016.0096
Moebius-Clune, B. N., D.J. Moebius-Clune, B.K. Gugino, O.J. Idowu, R.R. Schindelbeck, A.J. Ristow, H.M. van Es, J.E. Thies, H.A. Shayler, M.B. McBride, K.S.M Kurtz, D.W. Wolfe, & G.S. Abawi. (2017). *Comprehensive Assessment of Soil Health – The Cornell Framework* (3.2). Cornell University. http://soilhealth.cals.cornell.edu/training-manual/
North Dakota Agriculture Weather Network. (2019). *Corn Growing Degree Days (GDD)*. https://ndawn.ndsu.nodak.edu/help-corn-growing-degree-days.html
Nunes, M. R., Veum, K. S., Parker, P. A., Holan, S. H., Karlen, D. L., Amsili, J. P., Es, H. M., Wills, S. A., Seybold, C. A., & Moorman, T. B. (2021). The soil health assessment protocol and evaluation applied to soil organic C. *Soil Sci. Soc. Am. J.*, *85*, 1196–1213. https://doi.org/10.1002/saj2.20244
Oakridge National Laboratory. (2020). *DAYMET Daily Surface Weather and Climatological Summaries*. Single Pixel Extraction Tool. https://daymet.ornl.gov/single-pixel/
SAS Institute Inc. (2020). *SAS Programing Documentation*. https://documentation.sas.com/
Schindelbeck, R. R., Moebius-Clune, B. N., Moebius-Clune, D. J., Kurtz, K. S., & van Es, H. M. (2016). *Cornell University Comprehensive Assessment of Soil Health Laboratory Standard Operating Procedures*.
Soil Survey Staff. (2019). *Web Soil Survey*. Natural Resources Conservation Service, United States Department of Agriculture. http://websoilsurvey.sc.egov.usda.gov/
Stott, D. E. (2019). *Recommended Soil Health Indicators and Associated Laboratory Procedures. Soil Health Technical Note No. 450-03. U.S. Department of Agriculture, Natural Resources Conservation Service.* https://www.nrcs.usda.gov/wps/PA_NRCSConsumption/download?cid=nrcseprd1420229&ext=pdf
Stott, D. E., Andrews, S. S., Liebig, M. A., Wienhold, B. J., & Karlen, D. L. (2010). Evaluation of β-Glucosidase Activity as a Soil Quality Indicator for the Soil Management Assessment Framework. *Soil Science Society of America Journal*, *74*(1), 107–119. https://doi.org/10.2136/sssaj2009.0029

Tremblay, N., Bouroubi, Y. M., Bélec, C., Mullen, R. W., Kitchen, N. R., Thomason, W. E., Ebelhar, S., Mengel, D. B., Raun, W. R., Francis, D. D., Vories, E. D., & Ortiz-Monasterio, I. (2012). Corn response to nitrogen is influenced by soil texture and weather. *Agronomy Journal*, *104*(6), 1658–1671. https://doi.org/10.2134/agronj2012.0184

Wade, J., Culman, S. W., Logan, J. A. R., Poffenbarger, H., Demyan, M. S., Grove, J. H., Mallarino, A. P., McGrath, J. M., Ruark, M., & West, J. R. (2020). Improved soil biological health increases corn grain yield in N fertilized systems across the Corn Belt. *Scientific Reports*, *10*(1), 1–10. https://doi.org/10.1038/s41598-020-60987-3

Ward Laboratories. (2020). *Haney Test Interpretation Guide v1.0*. https://www.wardlab.com/wp-content/uploads/2020/06/Haney-Rev-1.0-Interpretation-Guide-PDF.pdf

Warncke, D., & Brown, J. R. (1998). Potassium and Other Basic Cations. In *Recommended Chemical Soil Test Procedures for the North Central Region Publication No. 221 (revised)* (pp. 31–33). University of Missouri Ag. Exp. Station.

Weidhuner, A., Hanauer, A., Krausz, R., Crittenden, S. J., Gage, K., & Sadeghpour, A. (2021). Tillage impacts on soil aggregation and aggregate-associated carbon and nitrogen after 49 years. *Soil and Tillage Research*, *208*(November 2020), 104878. https://doi.org/10.1016/j.still.2020.104878

Wienhold, B. J., Karlen, D. L., Andrews, S. S., & Stott, D. E. (2009). Protocol for indicator scoring in the soil management assessment framework (SMAF). *Renewable Agriculture and Food Systems*, *24*(04), 260–266. https://doi.org/10.1017/S1742170509990093

Wood, S. A., & Bowman, M. (2021). Large-scale farmer-led experiment demonstrates positive impact of cover crops on multiple soil health indicators. *Nature Food*, *2*, 97–103. https://www.nature.com/articles/s43016-021-00222-y

Yost, M. A., Veum, K. S., Kitchen, N. R., Sawyer, J. E., Camberato, J. J., Carter, P. R., Ferguson, R. B., Fernández, F. G., Franzen, D. W., Laboski, C. A., & Nafziger, E. D. (2018). Evaluation of the Haney Soil Health Tool for corn nitrogen recommendations across eight Midwest states. *Journal of Soil and Water Conservation*, *73*(5), 587–592. https://doi.org/10.2489/jswc.73.5.587

CHAPTER 4

SOIL HEALTH INDICATOR, ASSESSMENT SCORE, AND YIELD RESPONSE TO COVER CROPS[3]

Abstract: Farmer participatory research in soil health is crucial to evaluating soil conservation practices like cover crops. The Soil Health Partnership (SHP) was a large farmer-led network of on-farm trials assessing soil health. This research fills the need for wide-scale assessments of soil health indicators, scores, and crop yield from on-farm research with consistent methods across site-years. The effect of one to four years of cover crops on twelve soil health indicators, three soil health assessment composite scores, and two crop yields was evaluated using a mixed model analysis of covariance with repeated measures. Data was collected from 35 SHP sites, composed of 45 site-years from 2015 to 2019, that applied single or mixed species winter cover crops in corn (*Zea mays* L.) and soybean (*Glycine max* L.) rotations. The first soil health measurements were used as a covariate in the analysis. Soil microbial respiration (C mineralization) using 96-hr incubation was the only indicator to respond to cover crops. Additionally, the Comprehensive Assessment of Soil Health assessment composite score also responded to the treatment. The treatment did not interact with the baseline for any of the response variables. These results demonstrate to practitioners that soil respiration might be a helpful soil health indicator to monitor for soil health changes within one to four years following the adoption of conservation practices across the Midwestern U.S. The results

[3] Coauthors: Matt Yost, Maria Bowman, Kristen Veum, John Stevens; A manuscript prepared according to the Journal of Soil and Water Conservation style guide.

also suggest that the initial soil health values may not be significant within the first four years of cover crop use.

Key words: soil health—cover crops—Midwestern U.S.—on-farm research

The Soil Health Partnership (SHP) was a grower-led on-farm participatory research initiative of the National Corn Growers Association, which operated from 2014 to 2021. The SHP that brought together universities, collaborators from federal agencies, farmer groups, private companies, and environmental groups to promote the adoption of soil health practices and study their economic and ecological benefits and risks (Soil Health Partnership, 2021). The SHP worked with growers throughout much of the Midwestern US by launching randomized and replicated strip trial evaluations of growers' historical management versus a soil health-promoting practice, typically cover crops or reduced tillage. Growers provided general information about their agronomic practices and crop yield, and soil samples were regularly collected and analyzed for a suite of soil health indicators that are typically use in soil health assessments, such as the Comprehensive Assessment of Soil Health (CASH), Soil Management Assessment Framework (SMAF), and the Haney Soil Health Tool (HSHT). Since Karlen et al. (2017) described early lessons learned from the SHP on-farm soil health evaluations to readers of this journal, several papers have reported portions of SHP soil health research results. For example, details have been published on soil health spatial and temporal variation (Crookston et al., 2021), nutrient management (Flis & Bowman, 2021), and cover crop treatments (Wood & Bowman, 2021). This manuscript endeavors to follow-up these publications by providing complementary details on the outcomes of the on-farm trials to promote soil health and the statistical effects of cover crops on a suite of soil health indicators, their assessment scores, and corn (*Zea mays* L.) and soybean (*Glycine max* L.) yields of 35 fields added to the network between 2014 and 2018.

The SHP data set represents year-over-year sample and data collection at working farms from a large geographic area, allowing research results to have broad generalizability to non-SHP farms within the region. This research fills the need for wide-scale estimates of the effect of cover crops on soil health indicators, scores, and crop yield from on-farm research with consistent methods across site-years.

Materials and Methods

Soil Health Data Collection from On-Farm Evaluations. The methods and soil health indicator analyses used to conduct the Soil Health Partnership on-farm evaluations were previously described by Wood and Bowman (2021) and Crookston et al. (2021). Briefly, the on-farm trials evaluated the impacts of soil health-promoting practices on soil health indicators. The trials compared randomized strips of the growers' historical management to a soil health treatment that primarily included cover crops or reduced tillage; each replicated four times. In this analysis, 35 SHP sites that joined the network between 2014 and 2018 and their data from 2015 to 2019 were used. Grain yield data were collected annually from combine-mounted yield monitoring systems with corresponding global positioning system locations.

Sampling for soil nutrient analyses at SHP sites occurred approximately annually between 2014 and 2019. Soil samples were collected at SHP sites roughly every other year beginning in 2015. However, some sites were sampled on different schedules.

Samples for soil P, K, organic matter (OM), and pH were collected to the 0–5 cm and 5–15 cm depths. Samples for soil active carbon (ActC), 96-hour microbial respiration assay (Resp [96 hr]), autoclaved citrate extractable protein index (ACE), wet aggregate stability (WAS), available water capacity (AWC), 24-hr microbial respiration (Resp [24 hr]), water-extractable organic C (WEOC), and water-extractable organic N (WEON) remained in one segment (0–15 cm), composited across a strip, and packed in coolers with ice packs for expedited shipping to soil analysis laboratories. Soil analyses at Ward Laboratories, Inc. (Kearney, NE) included loss-on-ignition OM, pH, P, and K. Soil samples in coolers were sent to Cornell University Soil Health Laboratory (Ithaca, NY) for the CASH suite analysis. Additional samples were sent to the USDA-Agriculture Research Service Grassland Soil Water Research Laboratory in Temple, TX for the Haney Soil Health Tool (HSHT) suite analysis (Haney et al., 2018). The USDA facility received samples from 2014 to 2018. Ward Labs completed the HSHT analyses during 2018 from select locations only and in 2019 using the same methods as the USDA lab. A summary of soil test methods for each indicator is available in table 4−1. Results were reported for the 0 to 15 cm depth layer to ensure nutrient and soil health test results' comparability. The two depths for the soil nutrient analyses were combined using weighted averages. The results from individual soil samples were averaged within a strip for each nutrient and soil health test.

Soil texture for each soil health sample was measured at the Cornell Soil Health Laboratory and classified using the USDA-NRCS system. Sand, silt, and clay observations reported by the Cornel Lab for each strip were averaged over time and were assigned to coarse (sand, loamy sand, and sandy loam), medium (loam, silt loam, silt), or fine (sandy clay loam, sandy clay, clay loam, silty clay loam, silty clay) texture groups.

Cover Crop Management. The SHP aimed to establish cover crop trials across the Midwestern US to investigate the economic and environmental impact of that soil health-promoting practice at working farms. Therefore, cover crop management was determined by the individual farm operators according to the annual needs of the farm operation. This allowed practices to vary from one site to another across the SHP. This context also allowed the farm operators to explore and learn how cover cropping works for them. Cover crop management practices were classified into four categories, namely, the number of cover crop species used annually, planting time relative to cash crop harvest, planting methods, and cover crop termination timing relative to cash crop planting. In most site-years, SHP operators used a single cover crop species, planted after cash crop harvest, and planted the cover crop seed using a grain drill. Further, most operators terminated the cover crop more than or within two weeks of cash crop planting. Table 4−2 lists the percent of site-years employing certain practices over the period when data was collected from SHP sites. Additionally, 78% of the sites had two years of cover crops, 6% had only one year, 8% had three years, and 8% had four years of cover crops (figure 4−1).

Calculating Soil Health Assessment Scores. Soil health indicators for the CASH, SMAF, and HSHT were scored using methods described in Fine et al. (2017), Andrews et al. (2004), and Haney et al. (2018), respectively. The CASH scores were calculated using the SHP database as the scoring calibration data set rather than the Cornell University

Soil Health Lab database and are referred to as $CASH_{SHP}$ in this manuscript. The mean and standard deviation of the indicator observed values are included herein for comparison with other published values. In the CASH framework, OM, ActC, ACE, Resp (96 hr), AGG, and AWC were scored using the cumulative normal distribution function with the mean and standard deviation of three soil texture groups (Fine et al., 2017). There were approximately 11%, 65%, and 24% of the observations in the coarse, medium, and fine texture groups, respectively. Therefore, before calculating the indicator scores, distribution normality was evaluated for each indicator within each soil texture group using Q-Q plots and histograms. Following evaluation, AGG lab values for coarse and medium soils were log-transformed. Additionally, ACE values were square root transformed for medium texture soils. Soil pH scores were calculated according to threshold values available in Fine et al. (2017). According to the CASH manual (Moebius-Clune et al., 2016), nutrient element scores are based on local recommendations with threshold values. Therefore, CASH-scored macro and micronutrients were omitted from this analysis rather than formulating each nutrient's scoring thresholds by state or sub-state regions.

The SMAF scores were calculated using algorithms and logic statements parameterized by location-specific environment factors such as climate, soil suborder, field slope, and soil texture (Andrews et al. 2004). The OM loss on ignition was converted to soil organic carbon (SOC) using the conversion factor 0.58 described by Cambardella et al. (2001). The location-specific factors were identified using the USDA-NRCS Web Soil Survey area of interest tool (Soil Survey Staff, 2019) for each location. It is important to note that although the SMAF AGG score was originally parameterized from a different aggregate stability analysis method than the one used at the Cornell University Soil Health Lab, this study scored the Cornell AGG test results using the $CASH_{SHP}$ and the SMAF scoring approaches to evaluate the scores for all available indicators. The $CASH_{SHP}$ and SMAF AGG scores were then subjected to regression analysis to determine the agreement between the methods (see the Results and Discussion section). The HSHT scores were calculated according to Equation 4−S1 available in the Supplementary Material. The $CASH_{SHP}$ and SMAF composite scores were an unweighted average of the individual indicator scores.

Analytical Approach. The effect of cover crops on yield, soil health indicators, and scores was evaluated by accounting for sources of non-experimental variation while testing the hypothesis of no treatment effects (Stroup et al., 2018). Among analytical factors, there were six independent variables: cover crops treatment (T), the baseline measure taken by the SHP of the response variable (BL), year, site, and strip within site. Additionally, there were 30 response variables: yield of two cash crops, lab values of 12 soil health indicators, three soil health assessment composite scores, and 13 soil health indicator assessment scores, each evaluated in independent models. Furthermore, two versions of the model were tested to assess the effects of cover crops using broad and narrow inference (Dixon et al., 2019).

Data Preparation. The data set was first filtered to include sites that had annual rotations of corn and soybean. Next, sites were identified that had at least two years of soil health indicator data and yield data from the same crop in those two years. Additionally, those sites needed to have had at least one year of cover crop application.

All other SHP sites that did not trial cover crops were excluded from the data set. Furthermore, several sites were excluded where cover crops had been inadvertently applied to the control and the treatment strips in different years. Detailed cover crop management data available in the spring of 2021 following an exhaustive survey of management practices across the SHP was utilized for final site selections. Several sites were identified using the updated data and excluded because they had already applied cover crops before the first soil health samples were collected and subsequently had no baseline soil health measurement. Once the data set had been finalized, the response variables were evaluated for normality and transformed using the process described above, albeit not divided by soil texture groups. See table 4−S1 for transformations. Following data set preparation, histograms that included the mean and standard deviation were made of the strip-level distributions of the soil health indicator lab values, crop yields, and soil health assessment scores. The CORR procedure of SAS (SAS Institute Inc., 2020) was then used to calculate the Pearson's correlation coefficient for the soil health indicator lab values.

Model Development. A general linear mixed model was developed by first identifying hypothesized sources of variation by decomposing the system of environment, management, and experimental factors (Gezan & Carvalho, 2018). Variation in the experimental units was accounted for by utilizing the initial measure of the response variable (units depended on the variable). The presence of cover crops (no, yes) was considered as the primary fixed effect, while year, site, and the interactions of site by treatment or year by treatment were fixed or random effects depending on the broad or narrow inference model. The first model utilized a statistical design to draw narrow inferences with site as a fixed effect and year and its interaction with treatment as random effects. The second model was parameterized using site and year and their interactions with the treatment as random effects to make broad inferences across the region to a hypothetical population of farms. Strip nested within site was the subject of the repeated measures analysis in both the narrow and broad inference models.

While cover crop management is a significant contributor of variation among experimental units, given the number of different cover crop management practices used at the SHP sites (see table 4−2), creating coded variables to represent each practice to be utilized within the model was untenable. Therefore, cover crop management was considered an aspect of the site, which was treated as fixed or random in the narrow and broad inference models, respectively. The fully specified model tested the response of a single dependent variable to the factorial interactions of the baseline measure and cover crop treatment in the broad inference model and the baseline, site, and cover crop treatment in the narrow inference model.

Model Selection. The linear mixed model analysis of covariance with repeated measures was specified in the GLIMMIX procedure of SAS (SAS Institute Inc., 2020). The best-fitting covariance structure for each response variable was selected by testing eight covariance structures and identifying the covariance structure with the lowest corrected Akaike's information criterion (Hurvich & Tsai, 1989). The *solution* and *cl* options were used in the MODEL statement to request the regression parameter estimates and their confidence intervals. Finally, a reduced model was identified by iteratively removing the non-significant independent factors until only significant factors remained (Stroup et al., 2018) unless the non-significant main effect was part of a significant interaction. This process was independently repeated for each response variable. Subsequently, the final models of each indicator were evaluated for influential observations and outliers following the model selection that remained after a transformation. The REG procedure of SAS generated the studentized residual plots used to identify influential observations for individual response variables that crossed the Bonferroni correction for studentized deleted residuals (Kutner et al., 2004) threshold—which was calculated with a custom SAS macro—and were then removed from the data set.

Results and Discussion

Soil Health Indicators, Scores, and Yield. Soil health and crop production were assessed at 35 SHP sites across much of the Midwest using twelve soil health indicators common to the CASH, SMAF, and HSHT. The indicator observed values, composite scores, and corn and soybean yields were typical of those observed in the region (Crookston et al., 2021) (figures 4−2, 3, and 4). However, the soil health indicators were only moderately correlated to each other (r = ~0.1 to 0.7) (table 4−S1).

Analysis of Main and Interaction Effects. The first analysis used the combined experiments model for the narrow inference that considered the site a fixed effect to determine how the treatment interacted with conditions at the site level. The analysis detected main effect treatment differences in three soil health indicators (OM, active carbon, and 96-hr respiration) and one soil health composite score ($CASH_{SHP}$; the individual indicator scores were not tested in the narrow inference models due to few responses from the indicator observed values). However, because the treatment by site interaction was not significant, no further insights could be gathered from the narrow inference analysis. This result indicated that broad inference was justified because cover crops did not influence soil health indicators or assessment scores differently by site.

In the broad inference model that considered sites and years as random effects, cover crops did not significantly interact with the baseline measure in any indicator lab values or crop yield. The cover crop treatment significantly affected 96-hr microbial respiration and the $CASH_{SHP}$ composite score only (table 4−3). The 96-hr microbial respiration treatment means for cover crops, and the control were 0.47 and 0.44 mg CO_2 C g^{-1} soil. The standard error of the mean was 0.027 mg CO_2 C g^{-1} soil. The $CASH_{SHP}$ composite score treatment means were 59.2 and 55.9 for cover crops and the control, respectively, and the standard error of the mean was 1.3. The baseline measure significantly affected all response variables except WEON. The nearly universal effect of the baseline measure, but the lack of interaction with treatment suggests that within the first one to four years of cover crops, the initial conditions do not play a large role in how quickly soil health improves. This may further suggest that regardless of whether soil health indicator values are high or low, soil health improvement is unobserved primarily in the early stages of cover crop implementation. These results have similarly been reported where cover crop effects on the soil and crop yield were measured within four years following adoption (Chamberlain et al., 2020).

The most significant finding from the broad inference analysis is that the 96-hr microbial respiration was affected by one to four years of cover crops. This agrees with the soil microbiome response to cover crops compared to bare fallow evaluated with a meta-analysis conducted by Kim et al. (2020). These authors reported that increased CO_2

respiration is likely associated with increased cellulose decomposition in cover crops than bared fallow. Furthermore, a recent report by Franzluebbers et al. (2021) reported that multispecies mixed cover crops improved soil microbial respiration across 15 locations in North Carolina with only one to two years of treatments. More roots contributed to greater soil C fractions in cover crops compared to no cover in their analysis. However, there was no significant soil microbial respiration response in the present study to cover crops for the 24-hr respiration test (table 4−3). Notably, the 96-hour and 24-hour respiration test values had low correlation (r = 0.29; table 4−S3). This low correlation may be due to soil test methods that use different amounts of soil, incubation times, and detection methods (Haney et al., 2018; Schindelbeck et al., 2016). Moreover, for soil-test biological activity, Franzluebbers (2020) found that when soil-test subsamples are greater than or equal to 50g, there is increased precision of the estimated soil microbial respiration. The respiration test methods used for this manuscript utilized less than 50g of soil. There is a possibility that for the present study, the use of different labs to conduct the 24-hr respiration test introduced additional experimental error that may have led to a non-significant response to the cover crop treatment.

For the indicator scores, the $CASH_{SHP}$ score for 96-hr respiration was the only $CASH_{SHP}$ or SMAF indicator score responsive to cover crops (p <0.05). All other $CASH_{SHP}$ and SMAF indicator scores were responsive to the baseline measure only. Therefore, the indicator scores ANCOVA results are not shown. The positive response to the cover crop treatment in the 96-hour respiration $CASH_{SHP}$ score is expected. It supports previous evidence that scores reflect the response of the observed values to changes in management (van Es & Karlen, 2019). The $CASH_{SHP}$ composite score is an average of the $CASH_{SHP}$ indicators; thus, the positive response to cover crops from the composite score might be demonstrating sensitivity to a single indicator that was also responsive. This might suggest that the composite score will signify a difference in management practices even if all but one of the indicators were non-responsive. The implication is that individual indicator scores in the $CASH_{SHP}$ might be more reliable when evaluating soil health dynamics.

Wood & Bowman (2021) also analyzed SHP data with 78 sites and approximately 1500 observations. They reported that active C, aggregate stability, OM, and 96-hr respiration had a positive response to a cover crop by years of cover crop interaction; the main effects of those were also reported as responsive to the cover crops and years of cover crops. However, the estimated regression coefficient 95% confidence intervals crossed zero (the null hypothesis value) in the regression models for each of those indicators. This suggests that although the coefficients on the interaction term were positive, the actual value may have been neutral or a negative response. Furthermore, Wood and Bowman (2021) designated years of SHP participation as years of cover crop treatment. Yet, detailed cover crop management data became available after their manuscript was published (see Methods section herein), which revealed that the years of actual cover crop application did not match the numbers of years of participation at many sites. Thus, the updated analysis in the present study showed that only 96-hr respiration might be able to identify changes in soil health in the first four years of cover crop use in corn-soybean rotations of the US. Midwest.

It is crucial to note a benefit of participatory research is to explore dynamic responses to soil health-promoting practices. At the same time, growers make adaptive decisions based on their operation's immediate and long-term needs. Future studies of the legacy the SHP has left might also investigate the social benefits of participation in the network and the economic outcomes of these soil health-promoting practices.

Summary and Conclusions

These results demonstrate the influence of cover crop management and initial conditions on soil health outcomes. Specifically, the initial measure strongly determines the continued state of these soil health indicators, scores, and crop yield. Undoubtedly, inherent temporal dynamics from weather, crop, and soil management contribute to overall variation that might mask influences from cover crops that develop over time that were not detected from this analysis of short timespan on-farm trials. However, this analysis detected small changes in soil microbial respiration due to cover crops. This early response (one to four years of cover crops) in microbial activity might be indicative of long-term soil health improvement that can be realized throughout the Midwestern US. These results demonstrate to practitioners that soil respiration might be the indicator of choice to monitor for soil health changes within one to four years following the adoption of conservation practices.

Table and Figures

Table 4–1

Soil health indicator abbreviations, units, description, soil health assessment (SHA), and laboratory analysis methods used in this study.

Soil health indicator	Description	SHA	Analysis method	Citation
Organic matter (O.M.); soil organic carbon (SOC)	Carbon-based materials originating from living organisms	CASH, SMAF †	Calculated as weight lost from a soil sample on ignition. SOC was calculated by multiplying percent O.M. by 0.58.	Schindelbeck et al., 2016; Cambardella et al., 2001
Permanganate oxidizable carbon (active carbon) (ActC)	A measure of readily available organic carbon energy source for soil microbes.	CASH, SMAF	Photospectrometry analysis of oxidized potassium permanganate extractant.	Schindelbeck et al., 2016
Autoclaved citrate extractable soil protein index (ACE)	A measure of organically bound nitrogen. Microbial activity makes this organic matter fraction available for plant use.	CASH	High pressure and temperature extraction of citrate solution.	Schindelbeck et al., 2016
Soil microbial respiration 96-hour incubation (Resp (96 hr))	A measure of soil microbial metabolic activity.	CASH	Quantification of CO_2 gas trapped in solution evolved from re-wetted soil incubated 96 hours.	Schindelbeck et al., 2016
Soil microbial respiration 24-hour incubation (Resp (24 hr))	A short duration measure of soil microbial metabolic activity.	HSHT	Paper chromatography quantification of CO_2 gas evolved from re-wetted soil incubated 24 hours.	(Haney, 2020; Ward Laboratories, 2020)
Water-extractable organic carbon (WEOC)	A measure of readily available organic carbon energy source for soil microbes.	HSHT	Quantification of organic C extracted with water from a soil sample.	(Haney, 2020; Ward Laboratories, 2020)
Water-extractable organic nitrogen (WEON)	A measure of organically bound nitrogen. Considered as a "nutritional" source for microbes.	HSHT	Quantification of organic N extracted with water from a soil sample.	(Haney, 2020; Ward Laboratories, 2020)
pH	Affects the availability of nutrients and biological properties in the soil.	CASH, SMAF	Voltage meter calibrated to determine Hydrogen ion activity in soil solution.	Watson and Brown, 1998
Soil chemical nutrients: P, K	Soil nutrients are needed for healthy plant growth.	SMAF	Mehlich-III extractant method and quantified using inductively coupled atomic plasma spectroscopy.	Soil and Plant Analysis Council, 1999; (Warncke & Brown, 1998)
Available water capacity (AWC)	Soil water available for plant uptake.	CASH, SMAF	Amount of water extracted from a pulverized and sieved soil sample using a pressure chamber.	Schindelbeck et al., 2016
Soil wet aggregate stability (WAS)	The proportion of soil aggregates resistant to degradation following rain.	CASH, SMAF	Calculated from soil remaining on a 0.25 mm sieve following simulated rainfall.	Schindelbeck et al., 2016
Sand, silt and clay	Soil proportions of particle size 0.002—0.05 mm (silt) and less than 0.002 mm (clay).		Rapid 4-hour quantification of sand, silt, and clay from soil/water solution.	Schindelbeck et al., 2016

† Comprehensive Assessment of Soil Health (CASH), Soil Management Assessment Framework (SMAF), and the Haney Soil Health Tool (HSHT).

Table 4–2
Percent of 40 site-years using a particular cover crop management practice.

Cover Crop Management	Site-years (%)
Species mix	
Single	75
Two or three	25
Planting time	
Post cash crop	82.5
Inter-seeded	5
Over-seeded	12.5
Planting method	
Drilled	50
Row crop planter	2.5
High clearance seeder	5
Broadcast with incorporation	27.5
Broadcast without incorporation	5
Aerial seeded	10
Termination timing	
Winterkill	12.5
Terminated more than two weeks before cash crop planting	30
Terminated within two weeks of cash crop planting	32.5
Terminated after planting	20
Terminated at planting	5

Figure 4–1
The number of years of cover crops at 35 Soil Health Partnership locations in the Midwestern US.

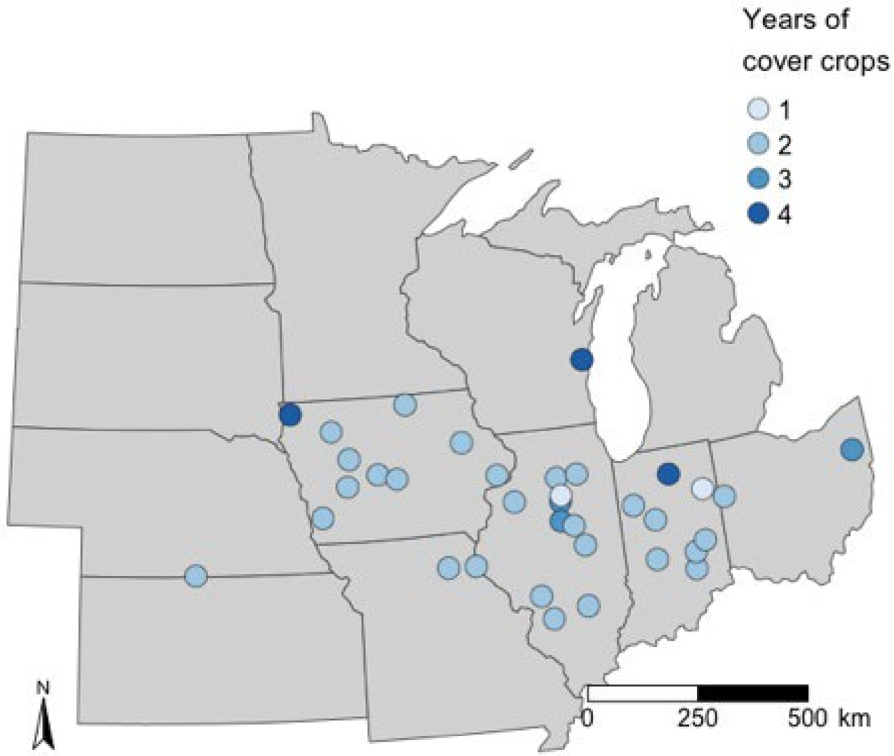

Figure 4–2

Soil health indicator strip-level observations (n = 362) at 35 Soil Health Partnership sites separated into biological (bio), chemical (chem), and physical (phys) categories. The mean is indicated by the dashed line; sd indicates the standard deviation. ACE, autoclaved citrate extractable protein index (mg g^{-1}); ActC, active carbon (mg kg^{-1}); AGG, wet aggregate stability (%); AWC, available water capacity (g g^{-1}); K, potassium (mg kg^{-1}); OM, organic matter loss on ignition (g kg^{-1}); P, phosphorus (mg kg^{-1}); Resp (24 hr), microbial respiration 24-hr incubation (mg CO_2 C kg^{-1}); Resp (96 hr), microbial respiration 96-hr incubation (mg CO_2 C g^{-1}); WEOC, water-extractable organic carbon (mg kg^{-1}); WEON, water-extractable organic nitrogen (mg kg^{-1}).

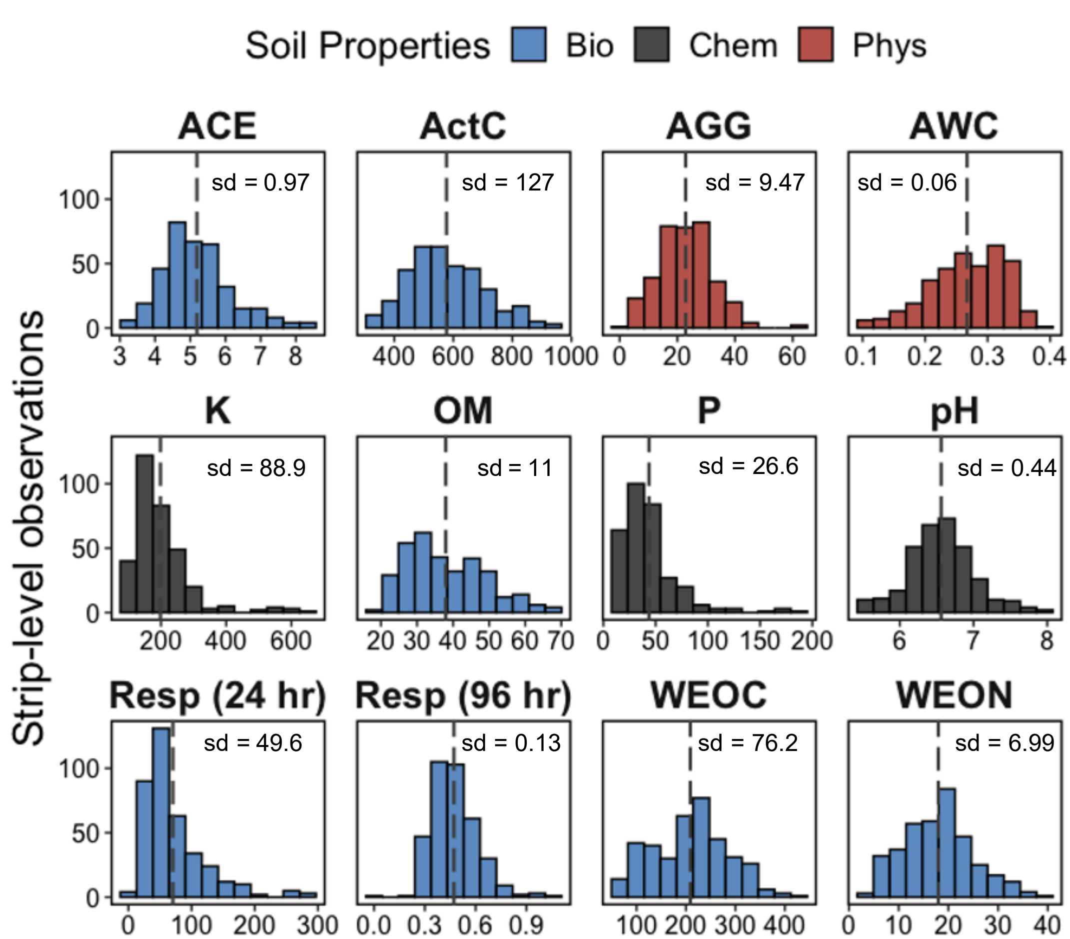

Figure 4–3

Corn and soybean yield strip-level observations (*n* = 148) at 35 Soil Health Partnership sites. The dashed line indicates the mean; sd indicates the standard deviation.

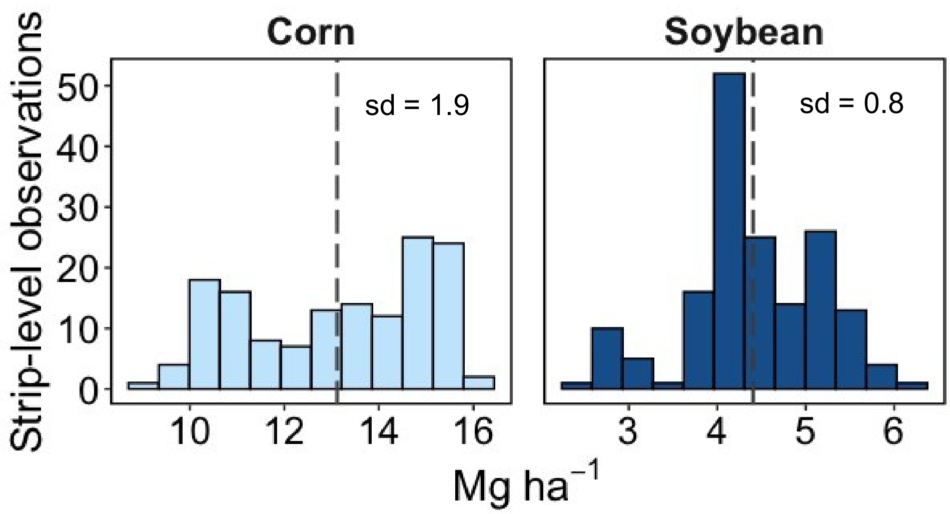

Figure 4–4

Strip-level observations at 35 Soil Health Partnership sites for the Comprehensive Assessment of Soil Health for the Soil Health Partnership (CASH SHP), Haney Soil Health Tool (HSHT), and the Soil Management Assessment Framework (SMAF). The dashed line indicates the mean; sd indicates the standard deviation. Sample sizes differed by assessment: $CASH_{SHP}$, *n* = 303; HSHT, *n* = 378; SMAF, *n* = 270.

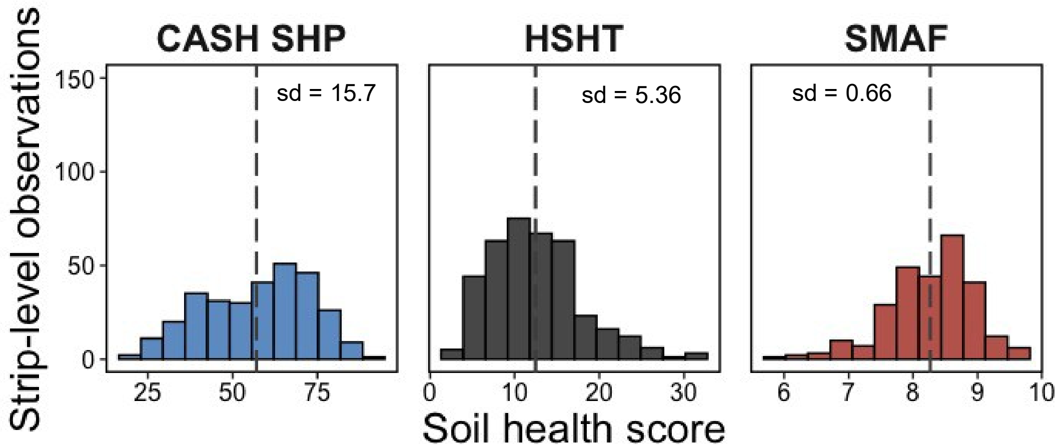

Table 4–3
Statistical significance from the analysis of covariance for cover crop treatment (T), the baseline measure of the response variable (BL), and their interaction. Rows represent independent models for the response variable and fixed effects. Non-significant fixed effects were iteratively removed from the individual models until only significant factors remained.

Response Variable	Fixed Effects		
	T	BL	T x BL
		P-value	
Soil health indicators			
OM		< 0.0001	
ActC		< 0.0001	
ACE		< 0.0001	
Resp 96 hr	0.0132	< 0.0001	
Resp 24 hr		< 0.0001	
WEOC		0.025	
WEON			
pH		< 0.0001	
P		< 0.0001	
K		< 0.0001	
WAS		< 0.0001	
AWC		< 0.0001	
Crops			
Corn		0.0018	
Soybean		< 0.0001	
Soil health assessment scores			
SMAF		< 0.0001	
CASH	0.0085	< 0.0001	
HSHT		< 0.0001	

Notes: ACE, autoclaved citrate extractable protein index; ActC, active carbon; AGG, wet aggregate stability; AWC, available water capacity; K, potassium; OM, organic matter loss on ignition; P, phosphorus; Resp (24 hr), microbial respiration 24-hour incubation; Resp (96 hr), microbial 0respiration 96-hour incubation; WEOC, water-extractable organic carbon; WEON, water-extractable organic nitrogen; SMAF, Soil Management Assessment Framework; CASH, Comprehensive Assessment of Soil Health; HSHT, Haney Soil Health Tool.

Supplementary Material

Equation 4–S1
Haney Soil Health Tool (HSHT) calculation:

When WEOC:WEON is < 5, then

$$HHMMHHHH = \frac{RRRRRRRR\ (24\ hr)}{10} + \frac{WWMMWWWW}{50} + \frac{WWMMWWWWW}{100}$$

If Resp (24 hr) < 100 mg CO_2 C kg^{-1} then,

$$HHMMHHHH = \frac{RRRRRRRR\ (24\ hr)}{10} + \frac{WWMMWWWW}{50} + \frac{WWMMWWWWW}{10}$$

if Resp (24 hr) is 100–200 mg CO_2 C kg^{-1} then,

$$HHMMHHHH = \frac{RRRRRRRR\ (24\ hr)}{12} + \frac{WWMMWWWW}{50} + \frac{WWMMWWWWW}{10}$$

if Resp (24 hr) is 200–300 mg CO_2 C kg^{-1} then,

$$HHMMHHHH = \frac{RRRRRRRR\ (24\ hr)}{14} + \frac{WWMMWWWW}{50} + \frac{WWMMWWWWW}{10}$$

if Resp (24 hr) is 300–400 mg CO_2 C kg^{-1} then,

$$HHMMHHHH = \frac{RRRRRRRR\ (24\ hr)}{16} + \frac{WWMMWWWW}{50} + \frac{WWMMWWWWW}{10}$$

if Resp (24 hr) is 400–500 mg CO_2 C kg^{-1} then,

$$HHMMHHHH = \frac{RRRRRRRR\ (24\ hr)}{18} + \frac{WWMMWWWW}{50} + \frac{WWMMWWWWW}{10}$$

if Resp (24 hr) is >500 mg CO_2 C kg^{-1} then,

$$HHMMHHHH = \frac{RRRRRRRR\ (24\ hr)}{20} + \frac{WWMMWWWW}{50} + \frac{WWMMWWWWW}{10}$$

where Resp (24 hr) is soil microbial respiration in 24-hour incubation, WEOC is water-extractable carbon, WEON is water-extractable nitrogen.

Table 4–S1

List of response variables and transformations used for data preparation.

Response Variable	Transformation
Soil health indicator observed values	
OM	Square root
ActC	None
ACE	Log
Resp 96 hr	Log
Resp 24 hr	Log
WEOC	Square root
WEON	Square root
pH	None
P	Log
K	Square root
WAS	Square root
AWC	Square
Crops	
Corn	None
Soybean	None
Soil health assessment composite scores	
SMAF	Square
CASH	None
HSHT	None

Notes: ACE, autoclaved citrate extractable protein index; ActC, active carbon; AGG, wet aggregate stability; AWC, available water capacity; K, potassium; OM, organic matter loss on ignition; P, phosphorus; Resp (24 hr), microbial respiration 24-hour incubation; Resp (96 hr), microbial respiration 96-hour incubation; WEOC, water-extractable organic carbon; WEON, water-extractable organic nitrogen.

Table 4–S2

Pearson correlation coefficients of 12 soil health indicators at 35 Soil Health Partnership. Statistical significance was determined at the 0.05 probability level.

	OM	**pH**	**P**	**K**	**AGG**	**AWC**	**ActC**	**ACE**	**Resp 96**	**Resp 24**	**WEOC**
pH	-0.11										
P											
K		-0.21	0.30								
AGG	0.34		-0.13	0.30							
AWC	0.20	-0.16	-0.13	0.16	-0.18						
ActC	0.77	0.11		0.36	0.34	0.23					
ACE	0.53	-0.22	0.25	0.48	0.35		0.53				
Resp 96	0.27	0.12			0.10		0.30	0.36			
Resp 24	0.40	-0.16		-0.11		0.17	0.22	0.31	0.29		
WEOC	0.46	-0.11	0.14				0.22	0.43	0.25	0.60	
WEON	0.29		0.23			-0.18		0.38	0.22	0.54	0.91

Notes: ACE, autoclaved citrate extractable protein index; ActC, active carbon; AGG, wet aggregate stability; AWC, available water capacity; K, potassium; OM, organic matter loss on ignition; P, phosphorus; Resp (24 hr), microbial respiration 24-hour incubation; Resp (96 hr), microbial respiration 96-hour incubation; WEOC, water-extractable organic carbon; WEON, water-extractable organic nitrogen.

Table 4–S3

Statistical significance from the analysis of covariance for cover crop treatment (T), site (S), the baseline measure of the response variable (BL), and their factorial interactions. Rows represent independent models for the response variable and fixed effects. Non-significant fixed effects were iteratively removed from the individual models until only significant factors remained.

Response Variable	Fixed Effects						
	T	S	BL	T x S	T x BL	S x BL	T x S x BL
	Statistical Significance†						
Soil health indicators							
OM‡	*	***	***			***	
ActC	*	***	***				
ACE		***	***				
Resp 96	**	***					
Resp 24		***	+				
WEOC		***					
WEON		***					
pH		**	***			**	
P		***	***			***	
K		***	***				
AGG		***	+				
AWC		***	***				
Crops							
Corn		***					
Soybean							
Soil health assessment scores							
SMAF		*	***			**	
CASH	***	*	***			+	
HSHT		*	+				

Notes: ACE, autoclaved citrate extractable protein index; ActC, active carbon; AGG, wet aggregate stability; AWC, available water capacity; K, potassium; OM, organic matter loss on ignition; P, phosphorus; Resp (24 hr), microbial respiration 24-hour incubation; Resp (96 hr), microbial respiration 96-hour incubation; WEOC, water-extractable organic carbon; WEON, water-extractable organic nitrogen; SMAF, Soil Management Assessment Framework; CASH, Comprehensive Assessment of Soil Health; HSHT, Haney Soil Health Tool.

† +Significant at the 0.1 probability level; *Significant at the 0.05 probability level; **Significant at the 0.01 probability level; ***Significant at the 0.001 probability level; ns, not significant.

‡ Due to missing values, only 34 sites were used in the OM analysis.

REFERENCES

Andrews, S. S., Karlen, D. L., & Cambardella, C. A. (2004). The Soil Management Assessment Framework. *Soil Sci. Soc. Am. J.*, *68*(6), 1945–1962. https://doi.org/10.2136/sssaj2004.1945

Cambardella, C. A., Gajda, A. M., Doran, J. W., Wienhold, B. J., & Kettler, T. A. (2001). Estimation of Particulate and Total Organic Matter by Weight Loss-on-Ignition. In L. R., K. J.M., F. R.F., & S. B.A. (Eds.), *Assessment Methods for Soil Carbon* (pp. 349–359). CRC Press.

Chamberlain, L. A., Bolton, M. L., Cox, M. S., Suen, G., Conley, S. P., & Ané, J. M. (2020). Crop rotation, but not cover crops, influenced soil bacterial community composition in a corn-soybean system in southern Wisconsin. *Applied Soil Ecology*, *154*(March), 103603. https://doi.org/10.1016/j.apsoil.2020.103603

Crookston, B. S., Yost, M. A., Bowman, M., Veum, K., Cardon, G., & Norton, J. (2021). Soil health spatial-temporal variation influence soil security on Midwestern, U.S. farms. *Soil Security*, *3*. https://doi.org/10.1016/j.soisec.2021.100005

Dixon, P. M., Moore, K. J., & van Santen, E. (2019). The Analysis of Combined Experiments. In B. Glaz & K. M. Yeater (Eds.), *Applied Statistics in Agriculture, Biological, and Environmental Science* (pp. 201–234). American Society of Agronomy, Crop Science Society of America, and Soil Science Socitey of America. https://doi.org/doi:10.2134/appliedstatistics.2016.0004

Fine, A. K., van Es, H. M., & Schindelbeck, R. R. (2017). Statistics, Scoring Functions, and Regional Analysis of a Comprehensive Soil Health Database. *Soil Sci. Soc. Am. J.*, *81*, 589–601. https://doi.org/10.2136/sssaj2016.09.0286

Flis, S., & Bowman, M. (2021). Soil Health and 4R: What Practices Are Working? *Crops & Soils*, *54*(1), 21–25. https://doi.org/10.1002/crso.20090

Franzluebbers, A. J. (2020). Soil mass and volume affect soil-test biological activity estimates. *Soil Science Society of America Journal*, *84*(2), 502–511. https://doi.org/10.1002/saj2.20038

Franzluebbers, A. J., Broome, S. W., Pritchett, K. L., Wagger, M. G., Lowder, N., Woodruff, S., & Lovejoy, M. (2021). Multispecies cover cropping promotes soil health in no-tillage cropping systems of North Carolina. *Journal of Soil and Water Conservation*, *76*(3), 263–275. https://doi.org/10.2489/jswc.2021.00087

Gezan, S. A., & Carvalho, M. (2018). Analysis of Repeated Measures for the Biological and Agricultural Sciences. In B. Glaz & K. M. Yeater (Eds.), *Applied Statistics in Agricultural, Biological, and Environmental Sciences* (pp. 279–298). American Society of Agronomy, Crop Science Society of America, and Soil Science Socitey of America. https://doi.org/10.2134/appliedstatistics.2016.0008

Haney, R. L., Haney, E. B., Smith, D. R., Harmel, R. D., & White, M. J. (2018). The soil health tool—Theory and initial broad-scale application. *Applied Soil Ecology*, *125*, 162–168. https://doi.org/10.1016/j.apsoil.2017.07.035

Hurvich, C. M., & Tsai, C.-L. (1989). Regression and Time Series Model Selection in Small Samples. *Biometrika*, *76*(2), 297–307. https://doi.org/https://doi.org/10.1093/biomet/76.2.297

Karlen, D. L., Goeser, N. J., Veum, K. S., & Yost, M. A. (2017). On-farm soil health evaluations: Challenges and opportunities. *Journal of Soil and Water Conservation*, *72*(2), 26A-31A. https://doi.org/10.2489/jswc.72.2.26A

Kim, N., Zabaloy, M. C., Guan, K., & Villamil, M. B. (2020). Do cover crops benefit soil microbiome? A meta-analysis of current research. *Soil Biology and Biochemistry*, *142*, 107701. https://doi.org/10.1016/j.soilbio.2019.107701

Kutner, M., Nachtsheim, C., & Neter, J. (2004). *Applied Linear Regression Model* (4th ed.). McGraw-Hill/Irwin.

Moebius-Clune, B. N., D.J. Moebius-Clune, B.K. Gugino, O.J. Idowu, R.R. Schindelbeck, A.J. Ristow, H.M. van Es, J.E. Thies, H.A. Shayler, M.B. McBride, K.S.M Kurtz, D.W. Wolfe, & G.S. Abawi. (2017). *Comprehensive Assessment of Soil Health – The Cornell Framework* (3.2). Cornell University. http://soilhealth.cals.cornell.edu/training-manual/

SAS Institute Inc. (2020). *SAS Programing Documentation*. https://documentation.sas.com/

Schindelbeck, R. R., Moebius-Clune, B. N., Moebius-Clune, D. J., Kurtz, K. S., & van Es, H. M. (2016). *Cornell University Comprehensive Assessment of Soil Health Laboratory Standard Operating Procedures*.

Soil Health Partnership. (2021). *About SHP*. Soil Health Partnership. https://www.soilhealthpartnership.org/we-are-shp/

Soil Survey Staff. (2019). *Web Soil Survey*. Natural Resources Conservation Service, United States Department of Agriculture. http://websoilsurvey.sc.egov.usda.gov/

Stroup, W. W., Miliken, G. A., Claassen, E. A., & Wolfinger, R. D. (2018). Analysis of Covariance. In *SAS for Mixed Models: Introduction and Basic Applications* (pp. 217–268).

van Es, H. M., & Karlen, D. L. (2019). Reanalysis Validates Soil Health Indicator Sensitivity and Correlation with Long-term Crop Yields. *Soil Science Society of America Journal*, *83*(3), 721–732. https://doi.org/10.2136/sssaj2018.09.0338

Wood, S. A., & Bowman, M. (2021). Large-scale farmer-led experiment demonstrates positive impact of cover crops on multiple soil health indicators. *Nature Food*, *2*, 97–103. https://www.nature.com/articles/s43016-021-00222-y

CHAPTER 5

CONCLUSION

These studies investigated how soil health relates to corn and soybean yield in three different ways. First, how variability in soil health affects variability in yield. Second, the strength of the relationships between soil health assessment scores—used to interpret soil health indicator observed values—and crop yield. And third, the effects of conservation management on soil health indicators, scores, and yield. These analyses found that soil health indicator temporal variation accounted for relatively little variability in corn and soybean yield over a two-to-four-year timespan. However, from the available data, the climate and soil tillage management had the most significant effect on corn and soybean yield variation over time.

Furthermore, soil health assessment composite scores were moderately correlated, meaning that the three assessments used tended to translate soil health in different ways, which may confuse how practitioners interpret soil health from one assessment to another. Additionally, soil health scores of individual indicators or composite scores were not often correlated with crop yield on a site-to-site basis. This might suggest to soil health researchers and growers that other soil health outcomes, such as field runoff water quality, be measured to determine how soil health is improving additional soil ecosystem services. Finally, the on-farm soil health trials revealed that few soil health indicators were affected by cover crops within a short one to four years of treatment timespan. Overall, these results suggest to growers that a whole-of-ecosystem approach be taken to monitoring soil health and that soil health measurements be taken before beginning a new conservation management plan, then every two to four years to allow time for soil health improvement.

APPENDIX

CURRICULUM VITAE

Education

Ph.D., Soil Science – Utah State University. Dissertation: *Soil Health Monitoring and Management in Corn-Soybean Agroecosystems of the Midwestern U.S.* (Dissertation successfully defended August 24, 2021)	Expected Dec. 2021
Graduate Certificate, Anticipatory Intelligence – Utah State University, Capstone: *The Art and Science of Anticipating the Future.*	Expected Dec. 2021
M.S., Plant, Soil, and Environmental Science – West Texas A&M University. Thesis: *Pearl Millet Forage Production and Water Use in the Texas High Plains*.	May 2018
B.S., University Studies – Brigham Young University-Idaho, Rexburg, ID.	Dec. 2015

Positions Held

Graduate Research Assistant, Utah State University, Logan UT.	2018 – Present
Temp. Instructor – PSES 2411 Soils and the Environment, West Texas A&M University, Canyon, TX.	2018–2018
Graduate Research Assistant, West Texas A&M University, Canyon, TX.	2016–2018

Publications

Refereed Journal Articles

Crookston, B.S., M.A. Yost, M. Bowman, K. Veum. Relationships of On-Farm Soil Health Scores with Corn and Soybean Yield in the Midwestern US. *Soil Science Society of America J.*	*In review*
Crookston, B.S., M.A. Yost, M. Bowman, K. Veum, G. Cardon, J. Norton. Soil Health Spatial-Temporal Variation Influence Soil Security on Midwestern, U.S. Farms. *Soil Security*, 3. *https://doi.org/10.1016/j.soisec.2021.100005*	2021
Crookston, B.S., B.C. Blaser, M. Darapuneni, M.B. Rhoades. Pearl Millet Water Use Efficiency. *Agronomy*, 10, 1672; doi:10.3390/agronomy10111672	2020

Extension Publications

Yost, M., G. Cardon, B. Crookston, J. Reeve, and E. Creech. Measuring and building soil health. Utah State Univ. Ext. Fact Sheet. 2018

Blog Posts

Crookston, B., and M. Yost. *New Article Highlights Variation in Soil Health Indicators Over Space and Time*. "Digging In" Blog. Soil Health Partnership. https://www.soilhealthpartnership.org/blog-story/new-article-highlights-variation-soil-health-indicators-over-space-time/ 2021

Crookston, B., and M. Yost. *Research Collaborations: SHP Supports Student*. "Digging In" Blog. Soil Health Partnership. https://www.soilhealthpartnership.org/blog-posts/research-collaborations-shp-supports-student. 2019

Videos

Crookston. B., and M. Yost. USU Crops Virtual Field Day. Logan UT. https://youtu.be/l1f08Yq5XwY 2020

Presentations

Guest Lectures

"Introduction to Complexity Science, Lecture 2: Modeling Complex Systems", CAI 5300 Art and Science of Anticipating the Future. Logan UT. Jan. 2020

"Introduction to Complexity Science, Lecture 1: Characteristics of Complex Systems", CAI 5300 Art and Science of Anticipating the Future. Logan UT. Jan. 2020

"Resilience in Agriculture," CAI 5100 Threats and Resilience. Logan UT. Mar. 2019

Professional Conferences

Oral Presentations

Crookston, B., M.A. Yost, M. Bowman, K. Veum, G. Cardon. Soil Health Indicator Temporal and Spatial Variation. ASA, CSSA, SSSA, annual conference, held virtually. 2020

Crookston, B., M.A. Yost. "Soil Health Testing." Soil and Water Conservation Training. Brigham City, UT. 2020

Crookston, B., M.A. Yost. Modeling relationships among indicators, management, environment, and yield for soil health planning. 2019. ASA, CSSA, SSSA Annual Meeting, San Antonio, TX. 2019

Crookston, B., M.A. Yost. 16. "Soil Health Partnership Analysis Updates." Scientific Advisory Committee of Soil Health Partnership, St. Louis, MO. 2019

Crookston, B., M.A. Yost, J. Cornell, K. Veum, D. Karlen, M. Bowman, and S. Mehta. 2019. The Soil Health Partnership database: Preliminary results and future prospects in soil health management. Soil Sci. Soc. Am. Annual 2019

Meeting, San Diego, CA.

Poster Presentations

Crookston, B., M. Yost, M. Bowman, J. Cornell, K. Veum, and S. Mehta. 2019. The Soil Health Partnership database: Preliminary results and future prospects in soil health management. Soil Health Institute Annual Meeting, Sacramento, CA. 2019

Crookston, B., and M. Yost. 2019. What are the themes and gaps in agriculture resilience literature? Plants, Soils, and Climate Department Showcase. Logan, UT. 2019

Crookston, B., M. Yost, J. Cornell, K. Veum, D. Karlen, and N. Goesser. 2018. The Soil Health Partnership Database: Goals and Prospects. Soil Health Institute Annual Meetings, Albuquerque, NM. 2018

Crookston, B., J. Machicek, B. Blaser, B.A. Stewart, M. Rhoades, M. Darapuneni. Water use of pearl millet forage in response to cultural practice in the semiarid Southern Great Plains. ASA, CSSA, SSSA Annual Meeting, 2017

Crookston, B., R. Spackman, P. Risen, Cutting management trial of quality-enhancing low lignin trait alfalfa. ASA, CSSA, SSSA Annual Meetings, Minneapolis, MN. 2015